

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 : Search time 21.7073 Seconds

(without alignments)
3044.503 Million cell updates/sec

Title: US-09-941-947a-18

Perfect score: 1623
Sequence: 1 MOYIYANRGFCAGVDRAIE.....GIEEKVVSIFPKLKKHQA 318

Scoring table: ELOSUM62
Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1623	100.0	318	US-09-934-903-24	Sequence 24, Appl
2	1623	100.0	318	US-09-941-847a-18	Sequence 18, Appl
3	1028.5	63.4	349	US-09-975-719-429	Sequence 429, Appl
4	787	48.5	338	US-10-156-761-10747	Sequence 10747, A
5	755	46.5	325	US-09-738-626-4643	Sequence 4643, Ap
6	743.5	45.8	329	US-09-712-363-271	Sequence 271, App
7	733	45.2	335	US-09-712-363-188	Sequence 188, App
8	663	40.9	310	US-10-289-762-10930	Sequence 10930, Ap
9	663	40.9	310	US-10-022-832-36	Sequence 36, Appl
10	598.5	36.9	539	US-10-275-360-5	Sequence 5, Appl
11	111	6.8	917	US-10-156-761-10047	Sequence 10047, A
12	109	6.7	578	US-09-712-363-282	Sequence 282, App
13	109	6.7	925	US-10-369-493-5415	Sequence 5415, App
14	104.5	6.4	501	US-10-369-493-1052	Sequence 1052, Ap
15	104.5	6.4	1130	US-10-369-493-19263	Sequence 19263, A

16	104	6.4	869	12	US-10-369-493-19983	Sequence 19983, A
17	103	6.3	433	9	US-09-815-242-10607	Sequence 10607, A
18	103	6.3	553	12	US-10-033-585-7338	Sequence 7338, Ap
19	102	6.3	1382	10	US-09-728-653-2	Sequence 2, Appl
20	101	6.2	391	12	US-10-369-493-20969	Sequence 20969, A
21	101	6.2	520	12	US-10-369-493-2893	Sequence 2893, Ap
22	100.5	6.2	538	12	US-10-369-493-2950	Sequence 2950, Ap
23	99.5	6.1	1765	15	US-10-037-182-8	Sequence 8, Appl
24	99.5	6.1	1786	9	US-09-873-676-113	Sequence 113, Appl
25	99.5	6.1	1786	15	US-09-938-275-6	Sequence 6, Appl
26	99.5	6.1	1786	15	US-10-037-182-6	Sequence 6, Appl
27	97.5	6.0	607	12	US-10-368-493-19473	Sequence 19473, A
28	97.5	6.0	921	10	US-09-738-626-5659	Sequence 5659, Ap
29	97	6.0	299	12	US-10-369-493-10700	Sequence 10700, A
30	96.5	5.9	321	9	US-09-815-242-13454	Sequence 13454, A
31	96.5	5.9	1881	12	US-10-033-585-7646	Sequence 7646, Ap
32	96	5.9	952	12	US-10-108-260A-3284	Sequence 3284, Ap
33	95.5	5.9	330	12	US-10-369-493-23413	Sequence 23413, A
34	95.5	5.9	977	12	US-10-342-136-2	Sequence 2, Appl
35	94.5	5.8	759	12	US-10-369-493-18629	Sequence 18629, A
36	94	5.8	496	15	US-10-215-224-7	Sequence 7, Appl
37	94	5.8	496	15	US-10-263-677-9	Sequence 9, Appl
38	94	5.8	573	10	US-09-847-6378-8	Sequence 8, Appl
39	93.5	5.8	547	12	US-10-369-493-7917	Sequence 7917, Ap
40	93	5.7	563	12	US-10-341-434-135	Sequence 135, Ap
41	93	5.7	573	10	US-09-828-574-1	Sequence 1, Appl
42	93	5.7	573	12	US-10-307-326-1	Sequence 1, Appl
43	93	5.7	573	12	US-10-116-275-141	Sequence 141, Appl
44	93	5.7	573	12	US-10-367-093-12	Sequence 12, Appl
45	93	5.7	573	15	US-10-046-649-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1	
US-09-934-903-24	
Sequence 24, Application US/099344903	
Patent No. US20020102690A1	
GENERAL INFORMATION:	
APPLICANT: Kofias, Mathias	
APPLICANT: Odum, J. Martin	
APPLICANT: Schenzel, Andreas J.	
APPLICANT: No. US20020102690A1ton, Kelley C.	
APPLICANT: Tomb, Jean-Francois	
APPLICANT: Rouviere, Pierre	
APPLICANT: Picoteglio, Stephen	
APPLICANT: Cheng, Qiong	
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production	
FILE REFERENCE: C11646 US NA	
CURRENT APPLICATION NUMBER: US/09/934,903	
CURRENT FILING DATE: 2001-08-22	
PRIOR APPLICATION NUMBER: 60/229,907	
PRIOR FILING DATE: September 1, 2001	
NUMBER OF SEQ ID NOS: 24	
SOFTWARE: Microsoft Office 97	
SEQ ID NO 24	
LENGTH: 318	
TYPE: PRT	
ORGANISM: Methylobionas 16a	
US-09-934-903-24	
Query Match	
Best Local Similarity 100.0%; Score 1623; DB 10; Length 318;	
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	
1 MOYIYANRGFCAGVDRAIEIVDAIEAFAPIVREYVHNRTVYVGLKQKGVFIEEL 50	
1 SDVPVGSYLISAGVKEVQQAEEERQLTVFDATCELVTKVMQVAKAKGCEVILLIG 120	
61 SDVPVGSYLISAGVKEVQQAEEERQLTVFDATCELVTKVMQVAKAKGCEVILLIG 120	

QY 121 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 180
 DB 121 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 180
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 240
 DB 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 240
 QY 241 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 300
 DB 241 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 300
 QY 301 EEKXVFSIPKELKXKMOA 318
 DB 301 EEKXVFSIPKELKXKMOA 318

RESULT 2
 US-09-941-947a-18
 ; Sequence 18, Application US/09941947A
 ; Publication No. US20030003528A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brzostowicz, Patricia C.
 ; APPLICANT: Cheng, Qiong
 ; APPLICANT: DiCosimo, Deana J.
 ; APPLICANT: Koffas, Vatcheos
 ; APPLICANT: Miller, Edward S. Jr.
 ; APPLICANT: Odom, J. Martin
 ; APPLICANT: Picataggio, Steve
 ; APPLICANT: Roviello, Pierre B.
 ; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
 ; FILE REFERENCE: CIL903 US NA
 ; CURRENT APPLICATION NUMBER: US/09/941.947A
 ; PRIOR FILING DATE: 2001-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,907
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,858
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 18
 ; LENGTH: 318
 ; TYPE: PRT
 ; ORGANISM: Methylobionas 16a
 US-09-941-947a-18

Query Match 100.0%; Score 1623; DB 11; Length 318;
 Best Local Similarity 100.0%; Pred. No. 2,5e-150;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOITLANPRGFCAGVDRAIEIVDOAIEAFGAPITYVREHVNRTVVYDGLKOKGAVFIEEL 60
 DB 1 MOITLANPRGFCAGVDRAIEIVDOAIEAFGAPITYVREHVNRTVVYDGLKOKGAVFIEEL 60
 QY 61 SDVPVGSYLIFSAGHVSKEVQOBAEROLTVFDPATCEPLTVTKVMQVAKHAKOGREVLILG 120
 DB 61 SDVPVGSYLIFSAGHVSKEVQOBAEROLTVFDPATCEPLTVTKVMQVAKHAKOGREVLILG 120
 QY 121 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 180
 DB 121 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 180
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 240
 DB 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 240
 QY 241 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 300
 DB 241 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 300
 QY 301 EEKXVFSIPKELKXKMOA 318
 DB 301 EEKXVFSIPKELKXKMOA 318

DB 301 EEKXVFSIPKELKXKMOA 318

RESULT 3
 US-09-975-719-429
 ; Sequence 429, Application US/09975719
 ; Publication No. US20030022349A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ausubel, Frederick M.
 ; APPLICANT: Rahme, Laurence G.
 ; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
 ; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
 ; FILE REFERENCE: 00786/361003
 ; CURRENT APPLICATION NUMBER: US/09/975.719
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 09/199,637
 ; PRIOR FILING DATE: 1998-11-25
 ; PRIOR APPLICATION NUMBER: US 60/066,517
 ; PRIOR FILING DATE: 1997-11-25
 ; NUMBER OF SEQ ID NOS: 437
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 429
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-975-719-429

Query Match 63.4%; Score 1028.5; DB 11; Length 349;
 Best Local Similarity 63.3%; Pred. No. 4,7e-92;
 Matches 198; Conservative 49; Mismatches 63; Indels 3; Gaps 2;

QY 1 MOITLANPRGFCAGVDRAIEIVDOAIEAFGAPITYVREHVNRTVVYDGLKOKGAVFIEEL 60
 DB 36 MOITLANPRGFCAGVDRAIEIVDOAIEAFGAPITYVREHVNRTVVYDGLKOKGAVFIEEL 95
 QY 61 SDVPVGSYLIFSAGHVSKEVQOBAEROLTVFDPATCEPLTVTKVMQVAKHAKOGREVLILG 120
 DB 96 SDVPVGSYLIFSAGHVSKEVQOBAEROLTVFDPATCEPLTVTKVMQVAKHAKOGREVLILG 155
 QY 121 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 180
 DB 156 HAGHEVEGTMGVEKCTEGGCIYLVETPEDVYENLKVNPNNDLAVYVOTTLSTDTKVMV 214
 QY 181 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 240
 DB 215 DALREOPPSIKEOKKDDICATONRQDAVHDLAKISDILLVYVGSPPNSNSNRLREIAVQL 274
 QY 241 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 300
 DB 275 GKPAYLIDTYODLKQDMLEGIYVGVGTAGASAPVLYVQVETIDOLKAMGGETTYSVRENSGI 332
 QY 301 EEKXVFSIPKELK 313
 DB 333 EEKXVFSIPKELK 345
 RESULT 4
 US-10-156-761-10747
 ; Sequence 10747, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 248-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; PRIOR FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10747
LENGTH: 338
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-10747

Query Match 48.5%; Score 787; DB 15; Length 338;
Best Local Similarity 48.1%; Pred. No. 2e-68;
Matches 153; Conservative 71; Mismatches 86; Indels 8; Gaps 3;

QY 2 QIVLNPFGCAGVRAIEIVDOAIEAFGAPIYVREHVNRRTVVDGLKQKGAVFIEELS 61
DB 11 RVLLAARPGYCAVDRAIVAVKLEKQYAPYVREHVNRRTVVDGLKQKGAVFIEERKA 70
QY 62 DVPVGSYLIFSAHGVSAKVEQOEAEERQITVFDATCPLVTVYHMQVAKAKQREVIILIGH 121
DB 71 EVPRGATVMSFAHGVSAKVEQOEAEERQITVFDATCPLVTVYHMQVAKAKQREVIILIGH 130
QY 122 AGHEVEGTMQYKCTEGGGIYVETPEDEVNRUKVNPNDLAVYTQTTLSTMTDKVMD 181
DB 131 EGHEVEGTMQYKCTEGGGIYVETPEDEVNRUKVNPNDLAVYTQTTLSTMTDKVMD 185
QY 182 ALRQPFPSIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIIVAVQ 241
DB 186 ALRQPFPSIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIIVAVQ 245
QY 242 -KRAYLDTYQDLKQMDLEGIEVGVYTAGASAPVYVQEVYDOLXKMGSETTSVRENSGI 300
DB 246 ARDYLVDYQDLKQMDLEGIEVGVYTAGASAPVYVQEVYDOLXKMGSETTSVRENSGI 303
QY 301 EEKVPFSIPKELKQKMOA 318
DB 304 EESITFSLPKELRDLRA 321

RESULT 5
US-09-738-626-4643
Sequence 4643, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MITOZUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4643
LENGTH: 325
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4643

Query Match 46.5%; Score 755; DB 10; Length 325;

Best Local Similarity 48.3%; Pred. No. 2.5e-65;
Matches 152; Conservative 67; Mismatches 86; Indels 10; Gaps 4;

QY 2 QIVLNPFGCAGVRAIEIVDOAIEAFGAPIYVREHVNRRTVVDGLKQKGAVFIEELS 61
DB 15 RVLLAARPGYCAVDRAIVAVKLEKQYAPYVREHVNRRTVVDGLKQKGAVFIEERKA 74
QY 62 DVPVGSYLIFSAHGVSAKVEQOEAEERQITVFDATCPLVTVYHMQVAKAKQREVIILIGH 121
DB 71 EVPRGATVMSFAHGVSAKVEQOEAEERQITVFDATCPLVTVYHMQVAKAKQREVIILIGH 134
QY 122 AGHEVEGTMQYKCTEGGGIYVETPEDEVNRUKVNPNDLAVYTQTTLSTMTDKVMD 179
DB 131 EGHEVEGTMQYKCTEGGGIYVETPEDEVNRUKVNPNDLAVYTQTTLSTMTDKVMD 189
QY 180 ALRQPFPSIKQKDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIIVAVQ 239
DB 190 VRELKVPQYQDPSPDDICVATONRODAVHDLAKISDILIVGSPNSNSNRLREIIVAVQ 249
QY 240 LGRP-AYLDTYQDLKQMDLEGIEVGVYTAGASAPVYVQEVYDOLXKMGSETTSVRENS 298
DB 250 NGADNAYLVYARIDPAMFEGETTIGISSGASVPELTVGVIERLAPGYD--DVEEVT 307
QY 299 GIEEKVPFSIPKELK 313
DB 308 SAEKIVPALPRVLR 322

RESULT 6
US-09-712-363-271
Sequence 271, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712, 363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179, 531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117, 844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118, 206
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126, 593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134, 093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134, 092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165, 124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165, 086
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 292
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 271
LENGTH: 329
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-712-363-271

Query Match 45.8%; Score 743.5; DB 10; Length 329;
Best Local Similarity 47.6%; Pred. No. 3.4e-64;
Matches 150; Conservative 60; Mismatches 94; Indels 11; Gaps 3;
QY 3 QIVLNPFGCAGVRAIEIVDOAIEAFGAPIYVREHVNRRTVVDGLKQKGAVFIEELS 62
DB 15 RVLLAARPGYCAVDRAIVAVKLEKQYAPYVREHVNRRTVVDGLKQKGAVFIEERKA 74

Db 20 VLLASPRSCAGSERALITYKRLVDVABGVYRKQIVANTVVAELRPGAVFVEDDE 79
 Qy 63 V---PVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEVQAKAKOGREYIL 118
 Db 80 IPBPFGAVVPSAIGVSPAVRAGADERGLQVVDATCPULVAKVHEAARFPAARDPTWIF 139
 Qy 119 IGHAGHEVGTMGQYKCEKTEGGGIVLVERPEDVRNKKNNPNDLAYVYQTLSTMDTKV 178
 Db 140 IGHAGHEVGTMGQYKCEKTEGGGIVLVERPEDVRNKKNNPNDLAYVYQTLSTMDTKV 194
 Qy 179 MDALRQPSISKEQKKDDI CYATONRQDAVHDLAKISDLIVGSPNSNSNRLRELAIV 238
 Db 195 VIDLRARFPFLGQPSIEDICVATNRQKRALQSMWECVYLVIGSCNSNSRRLRELAIV 254
 Qy 239 QLGKPAIDTYQDLKODMLGIEVVGVTAGASAPVLYQEVITDOLKAMGGETTSVRENSGI 298
 Db 255 RSGTPAYLIDGPPDIPEWMLSSVITGVITAGASAPRLVQVIDALUKYASIT--VYENS 312
 Qy 299 GIEKVFSPKELK 313
 Db 313 IATETVRFGLPKQVR 327

RESULT 7
 US-09-712-363-188
 ; Sequence 188, Application US/09712363
 ; Patent No. US20020164588A1

GENERAL INFORMATION:
 APPLICANT: Eisenberg, David
 APPLICANT: Rotstein, Sergio H.
 APPLICANT: Marcotte, Edward M.
 TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
 TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
 FILE REFERENCE: 07419-032001
 CURRENT APPLICATION NUMBER: US/09/712,363
 CURRENT FILING DATE: 2000-11-13
 PRIOR APPLICATION NUMBER: PCT/US00/02246
 PRIOR FILING DATE: 2000-01-28
 PRIOR APPLICATION NUMBER: 60/1179,531
 PRIOR FILING DATE: 2000-02-01
 PRIOR APPLICATION NUMBER: 60/117,844
 PRIOR FILING DATE: 1999-01-29
 PRIOR APPLICATION NUMBER: 60/118,206
 PRIOR FILING DATE: 1999-02-01
 PRIOR APPLICATION NUMBER: 60/126,593
 PRIOR FILING DATE: 1999-03-26
 PRIOR APPLICATION NUMBER: 60/134,093
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 60/134,092
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 60/165,124
 PRIOR FILING DATE: 1999-11-12
 PRIOR APPLICATION NUMBER: 60/165,086
 PRIOR FILING DATE: 1999-11-12
 NUMBER OF SEQ ID NOS: 292
 SOFTWARE: FASTSEQ for windows version 4.0
 SEQ ID NO 188
 LENGTH: 335
 TYPE: PRT
 ORGANISM: Mycobacterium tuberculosis
 US-09-712-363-188

Query Match 45.2%; Score 733; DB 10; Length 335;
 Best Local Similarity 47.0%; Pred. No. 3.8e-63;
 Matches 147; Conservative 66; Mismatches 92; Indels 8; Gaps 3;

Qy 2 QIVLANPFGCAGVRAIEIVDOAIEAFGAPIYVHEVNNRTVVDGLKQKAVFIEELS 61
 Db 27 RVLAEPRGVCAGVRAVETVERALQKHPYVYVHEIVNRRHVDTLAKAGAVFEETE 86
 Qy 62 DVPVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEVQAKAKOGREYILGH 121
 Db 87 QVPEGAIVVPSAIGVSPAVRAGADERGLQVVDATCPULVAKVHEAARFPAARDPTWIF 146

Qy 122 AGHPEVGTMGQYKCEKTEGGGIVLVERPEDVRNKKNNPNDLAYVYQTLSTMDTKVMD 181
 Db 147 EGHPEVGTMGQYKCEKTEGGGIVLVERPEDVRNKKNNPNDLAYVYQTLSTMDTKVMD 201
 Qy 182 ALRQPSISKEQKKDDI CYATONRQDAVHDLAKISDLIVGSPNSNSNRLRELAIVQJG 241
 Db 202 RLRRFPFLGQPSIEDICVATNRQKRALQSMWECVYLVIGSCNSNSRRLRELAIVQJG 261
 Qy 242 -KQAYLIDTYQDLKODMLGIEVVGVTAGASAPVLYQEVITDOLKAMGGETTSVRENSGI 300
 Db 262 ARALHVDVADDDIDSAHMDGVTTVGTSGASVPEVLYRQVRLRACGYDI--VQVPTTA 319
 Qy 301 EEKVFSPKELK 313
 Db 320 NETLVFALPRELR 332

RESULT 8
 US-10-289-762-1090
 ; Sequence 1090, Application US/10289762
 ; Publication No. US20040006218A1

GENERAL INFORMATION:
 APPLICANT: Griffls, R.
 TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragm
 TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, i
 TITLE OF INVENTION: and treatment of infection
 FILE REFERENCE: 9710-003-999
 CURRENT APPLICATION NUMBER: US/10/289,762
 CURRENT FILING DATE: 2003-03-27
 NUMBER OF SEQ ID NOS: 6849
 SEQ ID NO 1090
 LENGTH: 310
 TYPE: PRT
 ORGANISM: Chlamydia pneumoniae
 US-10-289-762-1090

Query Match 40.9%; Score 663; DB 12; Length 310;
 Best Local Similarity 45.0%; Pred. No. 2.4e-56;
 Matches 143; Conservative 59; Mismatches 96; Indels 18; Gaps 3;

Qy 2 QIVLANPFGCAGVRAIEIVDOAIEAFGAPIYVHEVNNRTVVDGLKQKAVFIEELS 61
 Db 3 KLILCNPRGSCGVRAIQVEVALIEKMGAPIYVKEHIVNRRHVNLRAKKAIFVEELV 62
 Qy 62 DVPVGSYLIFSANGSKEVQOAEEROLYFADATCPULVTKEVQAKAKOGREYILGH 121
 Db 63 DVEGERVYSAHGIPESVRAEARRKIIDATGSLVTKVSAKLYASKGYKILIGH 122
 Qy 122 AGHPEVGTMGQYKCEKTEGGGIVLVERPEDVRNKKNNPNDLAYVYQTLSTMDTKVMD 181
 Db 123 KKHVEVIGIVEVPE-----HITVEKVADEVALPFSDDPLFIYITQTLSDIVQEISS 177
 Qy 182 ALRQPSISKEQKKDDI CYATONRQDAVHDLAKISDLIVGSPNSNSNRLRELAIVQJG 241
 Db 178 ALIKRYSPIITLPSISICVATNRQKRALQSMWECVYLVIGSCNSNSRRLRELAIVQJG 237
 Qy 242 KQAYLIDTYQDLKODMLGIEVVGVTAGASAPVLYQEVITDOLKAMGGETTSVRENSGI 301
 Db 238 VADILNPNEDIDNIVVHSGDIATATAGASTPEVDVQACIRKL-----SSLIRGQVE 290
 Qy 302 -----EEKVFSPKELK 313
 Db 291 NDIPAVEDVVFQLPRELR 308

RESULT 9
 US-10-022-832-36
 ; Sequence 36, Application US/10022832
 ; Publication No. US20030059896A1
 ; GENERAL INFORMATION:
 APPLICANT: COUTURE, FRANCE
 APPLICANT: HAMEL, JOSEPH

```

; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; TITLE OF INVENTION: NOVEL CHLAMYDIA ANTIGENS AND CORRESPONDING DNA
; FILE REFERENCE: BIOVAC-15
; CURRENT APPLICATION NUMBER: US/10/022,632
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/256,941
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 36
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-10-022-832-36

```

```

Query Match      40.9%; Score 663; DB 15; Length 310;
Best Local Similarity 45.0%; Pred. No. 2,4e-56;
Matches 143; Conservative 59; Mismatches 98; Indels 18; Gaps 3;

```

```

QY 2 QIVLNPFGCAGVDRAIEIVDQAEAFGAPYVYHREYVNRIVVDGLKQKGAFFIELS 61
DB 3 KLILNPRGFCGSGVVALQVYVVALEKKGAEIYVHHEIVHRRVYVNLRAKGAIFVEELV 62
QY 62 DYPVGSYLIFSAHGVSKVQQAERQUTVFDATCPVTKVHMVQVAKHAKQREVIILGH 121
DB 63 DYPEGERIYSAHG:PPSPVRAKAKRKIIDIDATCGLVTKVSAKLYASKGVKILILGH 122
QY 122 AGHPEVEGTMGQYEKTEGGGIVLYETPEDEVNKKNNPNDAVAVTOTLSTMDTKVMD 181
DB 123 KKHVVICIGVEVPE-----HIVYERKADYBALPSSDPLFIITVTLISLDVQGIS 177
QY 182 ALREOFPSEIKOKKDDICATQNRDVAHDIAKISDLILVVGSSNSSNRLEIAVQIG 241
DB 178 ALKRYPSIILPSSISCAVTNRQKALRSVLSRVYVYVGVDSNSSNRLEIAVLRG 237
QY 242 KPAVLIIDYODLKQDMEGIEVGTAGASAPVLYQGVTDOLKAKWGETTSVRENGIE 301
DB 238 VPADLNPEDIDITVYHSGDIATAGASTPEDVQACIKL-----SSLIPGLQVE 230
QY 302 -----EKVPSIPKEIK 313
DB 291 NDIFAVEDVVFQLPKEIKR 308

```

```

RESULT 10
US-10-275-360-5
; Sequence 5, Application US/10275360
; Publication No. US20030115634A1
; GENERAL INFORMATION:
; APPLICANT: JOMAA, Hassan
; TITLE OF INVENTION: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHESIS PATH
; FILE REFERENCE: JOMAA-12 (PCT)
; CURRENT APPLICATION NUMBER: US/10/275,360
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: PCT/EP01/04537
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: DE10021688.9
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 5
; LENGTH: 539
; TYPE: PRT
; ORGANISM: plasmodium falciparum
; US-10-275-360-5

```

```

Query Match      36.9%; Score 598.5; DB 15; Length 539;
Best Local Similarity 43.6%; Pred. No. 1.2e-49;
Matches 123; Conservative 55; Mismatches 99; Indels 5; Gaps 1;
QY 3 IVLANPRGFCAGVDRAIEIVDQAEAFGAPYVYHREYVNRIVVDGLKQKGAFFIELSD 62

```

```

DB 226 LVILNPRGFCGSGVSAIETVSECLFLFPPPIYKHAKIYHNDIVCKLEKEGAIFTEIEND 285
QY 63 VPVGSYLIFSAHGVSKVQQAERQUTVFDATCPVTKVHMVQVAKHAKQREVIILGH 122
DB 286 VPDGHILLYSAHGISPQIRBAKQKLEIDATCPVTKVHMVQVAKHAKQREVIILGH 345
QY 123 GHPEVEGTMGQYEKTEGGGIVLYETPEDEVNKKNNPNDAVAVTOTLSTMDTKVMD 182
DB 346 NHVEVIGTYNAEPCIT-----HIVENVDVQKLNPNLKKLFYVQTTLISMDDCALLVQK 400
QY 183 LREOFPSEIKOKKDDICATQNRDVAHDIAKISDLILVVGSSNSSNRLEIAVQIG 242
DB 401 LKRPPEHIEITPSSISCAVTNRQKALRSVLSRVYVYVGVDSNSSNRLEIAVLRG 460
QY 243 PAVLIIDYODLKQDMEGIEVGTAGASAPVLYQGVTDOLKAKWGETTSVRENGIE 284
DB 461 PAVLANVHDDIQQILKNVKNIALTSASTEBQETQKFNILL 502

```

```

RESULT 11
US-10-156-761-10047
; Sequence 10047, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272657
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 10047
; LENGTH: 917
; TYPE: PRT
; ORGANISM: Streptomyces avermectilis
; US-10-156-761-10047

```

```

Query Match      6.8%; Score 111; DB 15; Length 917;
Best Local Similarity 20.7%; Pred. No. 0.14;
Matches 68; Conservative 50; Mismatches 97; Indels 114; Gaps 15;

```

```

QY 4 VIANPRGFCAGVDRAIEIVD---QAEAFGAPYVYHREYVNRIVVDGLKQKGAFFIEL 60
DB 632 IITNPKRAAEALQVYVREMDLRDYDLAIFG---YRHIDPENAEIRNG-KVK----- 678
QY 61 SDPVGSYLIFSAHGVSKVQQAERQUTVFDATCPVTKVH--MOVAKHAKQREVIIL 118
DB 679 --LPEGS-----BRLSPYPLVIVDELADLMAVAPR----- 709
QY 119 IGHAGHPYVGTMGQYEKTEGGGIVLY--ETB-DVRN--LVKNPNDAVAVTOTLIS 172
DB 710 -----DVEDAIVRITQOLAPAGIHLVLAQORPSVVVYGLIAVAPSRILAFATSS--S 759
QY 173 MTDTKVMD-----ALREOFPSEIKOK----- 194
DB 760 LADSRVILIDQGAERLIGKDGILFPMGANKPTMGOAFTEDEVEAVVQHCXQMAVVF 819
QY 195 KDDICATQNRDVAHDIAKISDLI-----LVVGSVNSNSSNRLEIAVQIGKAYIIFDT 245
DB 820 RDVAVVGTQKLEIDEDIGDLDLILQAAELVAVSTQFGSTSMQKRLVGFARAGRLMDL 879
QY 250 YODLKQDMEGIEVGTAGASAPVLYQ 278

```

DB 880 -----MESRNVGSGSKARDVYK 900

RESULT 12
US-09-712-363-282
Sequence 282, Application US/09712363
Patent No. US20020164586A1
GENERAL INFORMATION:
APPLICANT: Eisendberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712,363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206,
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086
NUMBER OF SEQ ID NOS: 292
SOFTWARE: PASTESEQ for Windows Version 4.0
SEQ ID NO 282
LENGTH: 578
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-712-363-282

Query Match 6.7%; Score 109; DB 10; Length 578;
Best Local Similarity 20.4%; Pred. No. 0.11;
Matches 78; Conservative 61; Mismatches 105; Indels 138; Gaps 18;
DB 6 ANRPGFC-----AGVDAIREIVDAIEAFAPIVYHREVVH 41
67 ANPGGVCESSCVSLAPNAPGSDIVVELDASHGVDDTRRLDR--AFVAPVQSRVYV-- 121
42 NRTVVDGKKQKGAVFIEBLSVPGSYLIFSAGVSK-----EVQGEAREEQLVFPDA 94
122 -----FIVDEAHVTTAGNALLKIVEEPEHLITFIPT 155
DB 95 TCP-----LVTKVH-----MQVAKAKQGRREVILIGHAGPEVEGTGQYE 135
156 TEPEKVLPTIRSRTHHPRFLPPTMRALLARICQ--EGVVDDAVYLV----- 205
136 KCTGGGGLYIVTEPD--VAKLVNNNDIAVYTOR---TSMPTDKTMVDALEQRP 188
206 -IRAGGG-----SPRDLTSLVDQLLAGRAD--THVYTRALGLGVTDVALIDAV---D 254
DB 189 SIKKQKODICVATQRQDAVHD-----LAKISLLIVGSPNSNS----- 230
255 ALAACDAALFGALIESVLDGCHDRPRPATDLERFLLIVVQSVPDASAGVADAPEDAL 314
DB 231 NRLREIVAGKPAVLIIDYQDLKQDMLEGIEVVGVTAASAPVIVQEVITDOL---KAM 287
315 DMRGQARIGRAT--LTRYAVAVQAGL-----GEMRGATAPRLLIEVCARILLPSAS 366
DB 286 GGETSVRENGSIEKVVFSIP 309
367 DAEALQWRERIEIRLMSIP 388

RESULT 13
US-10-369-493-5415
Sequence 5415, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5415
LENGTH: 925
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5415

Query Match 6.7%; Score 109; DB 12; Length 925;
Best Local Similarity 23.1%; Pred. No. 0.23;
Matches 77; Conservative 53; Mismatches 134; Indels 70; Gaps 17;
DB 3 IVLANPFGCGVDAIREIVDAIEAFAPIVYHREVVHNRVVDGKKQKGAVFIEBLS 62
143 MILNACKTFEYRAQAEIHDVIT---VPVYF---TPARLAVEBAQAGGLTVQLIN 196
DB 63 VPVGSYLIIFSAHGVSKVQGEAREEQLTVFD---ATCPVTVYHQAQAKHAGREVL 118
197 DGTAAAL---SHGIFRKEKIGEKPORLMVDDMAAKTTATYAFKLVKEKYEKOPKQTVL 253
DB 119 -IGHAGHPEVAGTQGVKCTEGGGLYLVETPEDEVNKLKN---NENDLAVYVOTILSM 174
254 GVG-----PDRITGGIE-MTNRLRDLIMFE--KNYKTKMTNTRRA-MTKFSKME 303
DB 175 DTKVMDALNEOPPSIEKQKDDICVATQRQDAVHDLAKISLLIVGSPNSNSNRLR 234
304 RLKQVLSANHEHFAQISABHDIDAKLVREDFNH---LISIMESRFGSP-TEOLRMA 359
DB 235 EIAVOLGKPAVLIIDYQDLKQDMLEGIEVVGVTAASAPV--LVQEV----- 281
360 QIPF-----DDIDQFVLMG-----AGTRVPVQEVITQKTIGTKETKELNT 400
DB 282 DOLKANGGETSVRENGS-----IEKVVFSI 308
401 DEAVAMGALLPQAHLSKPGKVPKPNIEKVIIPV 434

RESULT 14
US-10-369-493-1052
Sequence 1052, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 1052
 LENGTH: 501
 TYPE: PRT
 ORGANISM: Methanococcus jannaschii
 US-10-369-493-1052

Query Match 6.4%; Score 104.5; DB 12; Length 501;
 Best Local Similarity 20.5%; Pred No. 0.24;
 Matches 69; Conservative 51; Mismatches 134; Indels 83; Gaps 15;

QY 14 GVDRAIRIVDQAI---EAFGAPIYR-----HEVWNR-RIVVDGLK 50
 DB 133 GSEGAIEDIDBAIETABAIQFPVYVYKASGGGKMGMSVAYSKEBELKEVIEBARILKASAF 192
 QY 51 QKGAVFIEELSDVP--VSSYLIFSAHG-----VSKEVQOZAEEROLITPDCTLVYKVN 103
 DB 193 GDTVFIERKYLENPHIEIOLGDKHGNIIHSGDECSIGRRHKL-TEKPSPIKTE-- 249
 QY 104 MGVAKHAKOGREVLIGHGHEVEGTMGQYEKTEGGGIVLVETPEVDRLAKVNPNDL 163
 DB 250 ---ELRERKGBAIIKAGKINYSAGTV---EFLYENGNFYFLR-----NTRIOVEHT- 297
 QY 164 AYTDTLSMTDTKYAVDALREQPSIKEKKDICTATQ---NRQAVHDLAKISDLIL 220
 DB 298 --VTGQVTGIDLVKAMIKIAGBELITLKQSDVKIRGHAIKCRINAEPLNDPVPQPKIX 355
 QY 221 VVGSPPNSNSNRLREIAVQIGKPAVLIDTYODLKQDMLEGBIEVGVGTAGASAPVIVQEV 280
 DB 356 IYRSP-----GGPVRIIS-----GVYGAEIIP-YIDSM 384
 QY 281 IDQLXAWG--ETTSVRENSGIEEKVFSIPKELKH 315
 DB 385 IAKLITYGNSREELAPMRBALREYIIGVKTNIPRH 421

RESULT 15
 US-10-369-493-19263
 ; Sequence 19263, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 19263
 ; LENGTH: 1130
 ; TYPE: PRT
 ; ORGANISM: Myxococcus xanthus
 ; US-10-369-493-19263

Query Match 6.4%; Score 104.5; DB 12; Length 1130;
 Best Local Similarity 16.9%; Pred No. 0.86;
 Matches 70; Conservative 64; Mismatches 124; Indels 155; Gaps 13;

QY 38 EYVHRTVVDGLKQKGAVFIEELSDV-----DVGSYLI-----FSA 73
 DB 604 EYAEELADTINAMTQTLISFAQCVTDVARTVGVGKLGAGAVPVGAGTMDLTNNVILLA 663
 QY 74 HGVSKVQOQEA-----ERQITV-----FDATCPVYTKVM 104
 DB 664 NMLTDQVKNIAETVTAIVAKGLSKITVDKGEVLELKTINTMVDQLNSPAAYTVKAK 723
 QY 105 QVAKAKOGREVLIGHAGHPEVEGTMGQYEKTEGGGIVLVETPEVDV----- 153

DB 724 EYGTBGLK-----GQAEVRGSGWMDLTDNNVFAVNLITQVRGIVRVTAVAN 774
 QY 154 -----NLKVNPNDLAVTQTLISMTDT-----KAVDALR-----EOPPSIKE 192
 DB 775 GDLNQLTWDKAGELAAALADTINAMTQTLISFAQCVTDVARTVGVGKLGAGAEVPGVAG 834
 QY 193 QKRD---DICTATQRQDAVHDLAKIS-----DL-----ILVGSPPNSNSNR 232
 DB 835 TKKDLTNVNNLANLTLQVKNIAEVTIAVANGDLSSKKITVDKGEVLELKTINTMVDQ 894
 QY 233 LREIAVQIGKPAVLIDT-----YODLKQDMLEGBIEVGVGTAGASAPV----- 275
 DB 895 LKAFAEVTVRAKEVGTBGLGGQADVHDLGSMWMDLTDNNVILAGNLTDQVRNIAKYTT 954
 QY 276 -----LVGEVINDLKXAMGETTSVRENSGIEEKV 304
 DB 955 AVANGDLSSKITVSXKGEVLELKTINTMVDQLRFASEVTVRAKEVGTBGLK 1007

Search completed: January 29, 2004, 16:21:19
 Job time : 22.7073 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2004, 15:54:19 ; Search time 10.7171 Seconds

(without alignments)
3044.503 Million cell updates/sec

Title: US-09-941-947a-14

Perfect score: 806
1 MIRVGMGYDVRHNDGHI.....LGEGRKGIGIAVQAVLLER 157

Sequence: BLOSUM62
Gapop 10.0, Gapext 0.5

Scoring table: 789580 segs, 207824079 residues

Searched: Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubppa/PC01_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/1/pubppa/PC01_NEW_PUB.pep:*
- 7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	806	100.0	157	10	US-09-934-903-10	Sequence 10, App1
2	806	100.0	157	10	US-09-934-868-66	Sequence 66, App1
3	806	100.0	157	11	US-09-941-947a-14	Sequence 14, App1
4	496	61.5	170	15	US-10-174-410-12	Sequence 12, App1
5	360.5	44.7	223	12	US-10-859-194A-162	Sequence 162, App
6	251	31.1	174	15	US-10-156-761-11502	Sequence 11502, A
7	247.5	30.7	168	12	US-10-128-713A-10	Sequence 10, App1
8	213.5	26.5	150	10	US-09-128-625-6421	Sequence 6421, App
9	153	19.0	32	15	US-10-174-410-135	Sequence 135, App
10	153	19.0	199	12	US-10-289-762-585	Sequence 585, App
11	143	17.7	32	15	US-10-174-410-129	Sequence 129, App
12	138	17.1	32	15	US-10-174-410-133	Sequence 133, App
13	138	17.1	32	15	US-10-174-410-134	Sequence 134, App
14	138	17.1	32	15	US-10-174-410-136	Sequence 136, App
15	138	17.1	32	15	US-10-174-410-149	Sequence 149, App

15	133	16.5	32	15	US-10-174-410-125	Sequence 125, App
16	127	15.8	32	15	US-10-174-410-131	Sequence 131, App
17	126	15.6	32	15	US-10-174-410-127	Sequence 127, App
18	126	15.6	32	15	US-10-174-410-150	Sequence 150, App
19	125	15.5	32	15	US-10-174-410-128	Sequence 128, App
20	124	15.4	32	15	US-10-174-410-122	Sequence 122, App
21	124	15.4	32	15	US-10-174-410-140	Sequence 140, App
22	123	15.3	32	15	US-10-174-410-121	Sequence 121, App
23	123	15.3	32	15	US-10-174-410-124	Sequence 124, App
24	123	15.3	32	15	US-10-174-410-148	Sequence 148, App
25	119	14.8	32	15	US-10-174-410-113	Sequence 113, App
26	116	14.4	32	15	US-10-174-410-132	Sequence 132, App
27	116	14.4	32	15	US-10-174-410-153	Sequence 153, App
28	115	14.3	32	15	US-10-174-410-130	Sequence 130, App
29	115	14.3	32	15	US-10-174-410-138	Sequence 138, App
30	115	14.3	32	15	US-10-174-410-154	Sequence 154, App
31	114	14.1	32	15	US-10-174-410-139	Sequence 139, App
32	114	14.1	32	15	US-10-174-410-142	Sequence 142, App
33	114	14.1	32	15	US-10-174-410-143	Sequence 143, App
34	114	14.1	32	15	US-10-174-410-156	Sequence 156, App
35	112	13.9	32	15	US-10-174-410-151	Sequence 151, App
36	112	13.9	32	15	US-10-174-410-233	Sequence 233, App
37	110	13.6	32	15	US-10-174-410-123	Sequence 123, App
38	109	13.5	32	15	US-10-174-410-158	Sequence 158, App
39	109	13.5	32	15	US-10-174-410-229	Sequence 229, App
40	109	13.5	32	15	US-10-174-410-230	Sequence 230, App
41	108	13.4	32	15	US-10-174-410-189	Sequence 189, App
42	106	13.2	32	15	US-10-174-410-126	Sequence 126, App
43	106	13.2	32	15	US-10-174-410-146	Sequence 146, App
44	106	13.2	32	15	US-10-174-410-155	Sequence 155, App
45	106	13.2	32	15	US-10-174-410-173	Sequence 173, App

ALIGNMENTS

RESULT 1
US-09-934-903-10
Sequence 10, Application US/0934903
Patent No. US20020102690A1
GENERAL INFORMATION:
APPLICANT: Koffas, Mattheos
APPLICANT: Odum, J. Martin
APPLICANT: Schenkel, Andreas J.
APPLICANT: No. US20020102690A1ton, Kelley C.
APPLICANT: Tomb, Jean-Francois
APPLICANT: Bouviere, Pierre
APPLICANT: Picataggio, Stephen
APPLICANT: Cheng, Qiong
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
FILE REFERENCE: C1646 US NA
CURRENT APPLICATION NUMBER: US/09/934,903
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: September 1, 2001
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
LENGTH: 157
TYPE: PRT
ORGANISM: Methylobionas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by ORFs
US-09-934-903-10
Query Match 100.0%; Score 806; DB 10; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.5e-81;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MIRVGMGYDVRHNDGHIILGVKIPYEGLEAHSDDGVVAAADATLGAALDICK 60
DB 1 MIRVGMGYDVRHNDGHIILGVKIPYEGLEAHSDDGVVAAADATLGAALDICK 60
OY 61 HPPDIDNFGADSRVLLRHVYGVIVKEXKXVNAVDVTLIAQAPKMLPHVPGKFRANIAD 120

```

Db      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

Query Match      100.0%; Score 806; DB 10; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.5e-81;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
Db      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

Qy      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
Db      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

Qy      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
Db      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

RESULT 3
US-09-941-947a-14
; Sequence 14, Application US/09941947A
; Publication No. US2003003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: Diccasimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre B.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: C14903 US NA
; CURRENT FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14

```

```

; LENGTH: 157
; TYPE: PRT
; ORGANISM: Methylomonas 16a
US-09-941-947a-14

Query Match      100.0%; Score 806; DB 11; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.5e-81;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
Db      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

Qy      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
Db      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

Qy      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
Db      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

RESULT 4
US-10-174-410-12
; Sequence 12, Application US/10174410
; Publication No. US2003007313A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; FILE REFERENCE: 2C-METHYL-D-BRYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE RECPs
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1, 69, 109, 115
; OTHER INFORMATION: Xaa = Selenomethionine
US-10-174-410-12

Query Match      61.5%; Score 496; DB 15; Length 170;
Best Local Similarity 61.8%; Pred. No. 7.7e-47;
Matches 97; Conservative 21; Mismatches 39; Indels 0; Gaps 0;

Qy      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
      1 MIRVGMGYDVHRENDGDHIIILGGYKIPYEKGLAEHSDGDVVLHAAADATILGAALGDIK 60
Db      3 LIRIGHGPDVHAFEDPPLIIGVEVYHGPINSHSGDVALHLLTDAIIGAAALGDIK 62
      61 HFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 120
      63 LFPDTPNFKGADSRVLLRHVYGIYKEKGYKLVNADVTIIAQAAPKMLPHVPGKRAIIAAD 122
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      123 LQCDIEQVNVLATTTTEKLGFTGDEGIAACEAVALLIR 159

Qy      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
Db      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157
      121 LETDVPFINVKATTTTEKLGFEGRKEGIAVQAVVLIIR 157

RESULT 5
US-10-259-194A-162
; Sequence 162, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.

```

```

; APPLICANT: Chassemian, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiaki
; APPLICANT: Krepes, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: Patentlist.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 162
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-259-194A-162

```

```

Query Match          44.7%; Score 360.5; DB 12; Length 223;
Best Local Similarity 46.8%; Pred. No. 1,1e-31;
Matches 73; Conservative 27; Mismatches 55; Indels 1; Gaps 1;

```

```

QY 3 RVGNGYDVRFPNDGDHIIIGVKKIPEYKGLFASDGDVYTHALADALGAAAGDICKF 62
DB 67 RVGNGYDVRFPNDGDHIIIGVKKIPEYKGLFASDGDVYTHALADALGAAAGDICKF 126
QY 63 PD-TDPNFKGADSVLRHYVGIYKKGKYLVAADVTTIIQAAPKMLPHVPGKRAMIAADL 121
DB 127 PDLDPNFKGADSVLRHYVGIYKKGKYLVAADVTTIIQAAPKMLPHVPGKRAMIAADL 186
QY 122 ETDVDFINVKATTTKLGFGKREGIAVQAVVLI 157
DB 187 GADPSVYVTKAKTEKYEKYSIGENRSIAAHTVVLIR 222

```

RESULT 6

```

US-10-156-761-11502
; Sequence 11502, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11502
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Streptomyces avermectilis
US-10-156-761-11502

```

```

Query Match          31.1%; Score 251; DB 15; Length 174;
Best Local Similarity 37.6%; Pred. No. 1e-19;
Matches 59; Conservative 23; Mismatches 67; Indels 8; Gaps 3;

```

```

QY 4 VGMGYDVRFPNDGDHIIIGVKKIPEYKGLFASDGDVYTHALADALGAAAGDICKF 62
DB 21 VGIQDTHAFEDGRLMCAKLGKGEGLAGSDADVVAHMAACNALFSAAGLDLQHP 80
QY 63 PDLDPNFKGADSVLRHYVGIYKKGKYLVAADVTTIIQAAPKMLPHVPGKRAMIAADL 122
DB 81 GIGREMGAGAGVYLLTFAATVYAAAGFTTIGNVAVQVGRPKL-----GRRDEAGKIL 135
QY 123 TDV--DFINVKATTTKLGFGKREGIAVQAVVLI 157
DB 136 SDVAGAPVSVGATTTDGLGFGKREGIAVQAVVLI 172

```

RESULT 7

```

US-10-128-713A-10
; Sequence 10, Application US/10128713A
; Publication No. US20030170847A1
; GENERAL INFORMATION:
; APPLICANT: Brannucci, Michael G
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production
; FILE REFERENCE: C-1788
; CURRENT APPLICATION NUMBER: US/10/128,713A
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis
US-10-128-713A-10

```

```

Query Match          30.7%; Score 247.5; DB 12; Length 158;
Best Local Similarity 37.8%; Pred. No. 2.2e-19;
Matches 59; Conservative 24; Mismatches 66; Indels 7; Gaps 2;

```

```

QY 2 IRVNGYDVRFPNDGDHIIIGVKKIPEYKGLFASDGDVYTHALADALGAAAGDICKF 61
DB 1 MRVGLGTDVHPDIEVGRCDMAAGLFEEDACSGSDVDVVAHLCALISAAGIGDGSV 60
QY 62 PDLDPNFKGADSVLRHYVGIYKKGKYLVAADVTTIIQAAPKMLPHVPGKRAMIAADL 121
DB 61 FGTGRFPNDGVSAGAMALFRRLLRBNQFTVGNAAVYTGRRKTL-----GPRRDEAGK 115
QY 122 ETDV--DFINVKATTTKLGFGKREGIAVQAVVLI 155
DB 116 LSDILGAPVSVSATTDTGLGIGTREGIAAATVLI 151

```

RESULT 8

```

US-09-738-626-6421
; Sequence 6421, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16

```

```

; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 6421
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6421

```

```

Query Match      26.5%; Score 213.5; DB 10; Length 160;
Best Local Similarity 32.9%; Pred. No. 1.3e-15;
Matches 51; Conservative 25; Mismatches 72; Indels 7; Gaps 2;

```

```

QY 3 RVGAGYDVHFRNDGHIILGGVKIPYKGLAHSDDGVVLAALADALIGAAALGDIKHF 62
DB 8 RVGIATDAHQIENAGKPCWIALCLPEGYDGGGSHDGVVAHAIVDALISASGLDGSFV 67
QY 63 PDIPTNFGADSRVLRVYGVYKGYKLVNADVTIIAQPRLPHVPCRMATIAADLE 122
DB 68 GVGREYDVGSGTOLKEVRELLSAGVYIGVNAAGLVGQTPKRF-----GPRREKQCVI 122
QY 123 TDV--DFINVKATTEKLGFEGRKXGIAVQAVALI 155
DB 123 SEIGAPCSLATTITDMNGFTGRSEGRASVATNAV 157

```

```

RESULT 9
US-10-174-410-135
; Sequence 135, Application US/10174410
; Publication No. US20030072134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs
; FILE REFERENCE: 524982000300
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 135
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-174-410-135

```

```

Query Match      19.0%; Score 153; DB 15; Length 32;
Best Local Similarity 87.5%; Pred. No. 7e-10;
Matches 28; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

```

```

QY 40 VLAHALADALIGAAALGDIKHFPTDPPNFKG 71
DB 1 VLAHALADALIGAAALGDIKHFPTDPPNFKG 32

```

```

RESULT 10
US-10-289-762-585
; Sequence 585, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffla, R.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762

```

```

; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO: 585
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-585

```

```

Query Match      19.0%; Score 153; DB 12; Length 199;
Best Local Similarity 29.4%; Pred. No. 8.5e-09;
Matches 47; Conservative 26; Mismatches 75; Indels 12; Gaps 5;

```

```

QY 3 RVGAGYDVHFRNDGHIILGGVKIPYKGLAHSDDGVVLAALADALIGAAALGDIKHF 56
DB 40 RTGIGQDSHRLPSSSTYPCILGSIITPDHCGFOANSDDIIFHAICNAISVTNKLILG 99
QY 57 DIGKHFPTDPPNFKG-ADSRVLRVYGVYKGYKLVNADVTIIAQPRLPHVPCRMATIAADLE 115
DB 100 KVADELQOT---RGITDSGIYLEALKSLK-PNOKISHVALITEGSRPFLCTLSALRQ 154
QY 116 NIADLETVDFINVKATTEKLGFEGRKXGIAVQAVALI 155
DB 155 NIAQVMNLTPTDITGTAISERGLSDFGGSGVOCFVLIV 194

```

```

RESULT 11
US-10-174-410-129
; Sequence 129, Application US/10174410
; Publication No. US20030072134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs
; FILE REFERENCE: 524982000300
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 129
; LENGTH: 32
; TYPE: PRT
; ORGANISM: C. perfringens
US-10-174-410-129

```

```

Query Match      17.7%; Score 143; DB 15; Length 32;
Best Local Similarity 78.1%; Pred. No. 8.9e-09;
Matches 25; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 40 VLAHALADALIGAAALGDIKHFPTDPPNFKG 71
DB 1 VLAHALADALIGAAALGDIKHFPTDPPNFKG 32

```

```

RESULT 12
US-10-174-410-133
; Sequence 133, Application US/10174410
; Publication No. US20030072134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; TITLE OF INVENTION: 2C-METHYL-D-ERYTHRITOL, 2,4-CYCLODIPHOSPHATE SYNTHASE MECPs
; FILE REFERENCE: 524982000300
; CURRENT APPLICATION NUMBER: US/10/174,410
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/299,058

```

```

; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 32
; TYPE: PRF
; ORGANISM: S. typhimurium
US-10-174-410-133

Query Match
Best Local Similarity 17.1%; Score 138; DB 15; Length 32;
Best Local Similarity 84.4%; Pred. No. 3.2e-08;
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

RESULT 13
US-10-174-410-134
; Sequence 134, Application US/10174410
; Publication No. US20030073134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; FILE REFERENCE: 524982000300
; CURRENT FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 32
; TYPE: PRF
; ORGANISM: S. enterica
US-10-174-410-134

Query Match
Best Local Similarity 17.1%; Score 138; DB 15; Length 32;
Best Local Similarity 84.4%; Pred. No. 3.2e-08;
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

RESULT 14
US-10-174-410-136
; Sequence 136, Application US/10174410
; Publication No. US20030073134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; FILE REFERENCE: 524982000300
; CURRENT FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 32
; TYPE: PRF
; ORGANISM: Escherichia coli

```

```

US-10-174-410-136

Query Match
Best Local Similarity 17.1%; Score 138; DB 15; Length 32;
Best Local Similarity 84.4%; Pred. No. 3.2e-08;
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

RESULT 15
US-10-174-410-149
; Sequence 149, Application US/10174410
; Publication No. US20030073134A1
; GENERAL INFORMATION:
; APPLICANT: Louie, Gordon V.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF
; FILE REFERENCE: 524982000300
; CURRENT FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 336
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 32
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-174-410-149

Query Match
Best Local Similarity 17.1%; Score 138; DB 15; Length 32;
Best Local Similarity 84.4%; Pred. No. 3.2e-08;
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 40 VTHALADAILGAAALGDIKGFPTDPNFKG 71
DB 1 VALHALTDALGAAALGDIKGFPTDPAFKG 32

Search completed: January 29, 2004, 16:21:17
Job time : 10.7171 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: January 29, 2004, 15:54:19 ; Search time 37.1146 Seconds
(without alignments)
3044.503 Million cell updates/sec

Title: US-09-941-947a-16
Perfect score: 2788
Sequence: 1 MTRKFFITGVSSLSGKGLA.....HALPSGFVEAAAKTKTGCTA 544

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2788	100.0	544	10	US-09-934-903-12
2	2788	100.0	544	10	US-09-934-868-70
3	2788	100.0	544	11	US-09-941-947a-16
4	2008	72.0	542	9	US-09-815-242-11933
5	1816	68.7	545	9	US-09-815-242-10275
6	1908.5	68.0	545	9	US-09-815-242-11175
7	1783	64.0	511	9	US-09-815-242-13865
8	1576	56.5	536	9	US-09-815-242-12372
9	1576	56.5	536	9	US-09-815-242-12978
10	1569	56.3	536	9	US-09-815-242-4984
11	1569	56.3	536	9	US-09-815-242-10727
12	1561	56.0	535	9	US-09-815-242-13376
13	1444.5	51.8	538	9	US-09-815-242-11320
14	1432.5	51.4	549	15	US-10-156-761-14031
15	1372	49.2	554	10	US-09-738-626-5066

16	1355.5	48.6	537	12	US-10-289-762-246	Sequence 246, App
17	1006.5	36.1	1003	12	US-10-161-051-91	Sequence 91, App1
18	513	18.4	298	12	US-10-264-049-2881	Sequence 2881, Ap
19	353.5	12.7	199	15	US-10-106-698-4980	Sequence 4980, Ap
20	345.5	12.4	195	9	US-09-925-297-773	Sequence 773, App
21	327.5	11.7	94	12	US-09-864-408A-3914	Sequence 3914, Ap
22	245	8.8	59	12	US-09-864-408A-2350	Sequence 2350, Ap
23	204	7.3	135	12	US-10-264-049-4217	Sequence 4217, Ap
24	126.5	4.5	238	12	US-10-314-657-34	Sequence 34, App1
25	126	4.5	374	12	US-10-369-493-10131	Sequence 10131, A
26	125.5	4.5	1390	12	US-10-369-493-11073	Sequence 11073, A
27	124	4.4	374	12	US-10-369-493-15844	Sequence 15844, A
28	124	4.4	374	12	US-10-369-493-16223	Sequence 16223, A
29	123.5	4.4	371	12	US-10-369-493-1338	Sequence 1338, Ap
30	122	4.4	845	10	US-09-738-626-5347	Sequence 5347, Ap
31	120	4.3	382	12	US-10-369-493-691	Sequence 691, App
32	120	4.3	382	12	US-10-369-493-15474	Sequence 15474, A
33	120	4.3	382	16	US-10-210-115-33	Sequence 33, App1
34	116.5	4.2	656	12	US-10-369-493-18670	Sequence 18670, A
35	115.5	4.1	650	12	US-10-369-493-1338	Sequence 1338, Ap
36	115.5	4.1	650	12	US-10-369-493-20401	Sequence 20401, A
37	115.5	4.1	1687	14	US-10-094-679-3	Sequence 3, App1
38	114	4.1	1071	12	US-10-369-493-23117	Sequence 23117, A
39	113.5	4.1	106	12	US-09-864-408A-2256	Sequence 2256, Ap
40	113	4.1	362	12	US-10-369-493-17332	Sequence 17332, A
41	112.5	4.0	371	12	US-10-369-493-11567	Sequence 11567, A
42	112.5	4.0	537	12	US-10-369-493-89	Sequence 89, App1
43	111.5	4.0	728	12	US-10-369-493-15429	Sequence 15429, A
44	111.5	4.0	728	12	US-10-369-493-15797	Sequence 15797, A
45	111.5	4.0	728	12	US-10-369-493-16179	Sequence 16179, A

ALIGNMENTS

RESULT 1
US-09-934-903-12
Sequence 12, Application US/09934903
Patient No. US20020102690A1
GENERAL INFORMATION:
APPLICANT: Kofias, Mattheos
APPLICANT: Odum, J. Martin
APPLICANT: Schenzle, Andreas J.
APPLICANT: No. US20020102690A1on, Kelley C.
APPLICANT: Rouviere, Pierre
APPLICANT: Picataggio, Stephen
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
FILE REFERENCE: C14646 US NA
CURRENT APPLICATION NUMBER: US/09/934,903
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,907
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 544
TYPES: PRT
ORGANISM: Methylobionas 16a
FEATURES:
US-09-934-903-12
OTHER INFORMATION: Amino acid sequences encoded by ORF6
Query Match 100.0%; Score 2788; DB 10; Length 544;
Best local similarity 100.0%; Pred. No. 1.5e-254;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTRKFFITGVSSLSGKGLAASSLAIEDRGKVTITTKDDPYINVDGMSPPQGEVF 60
DB 1 MTRKFFITGVSSLSGKGLAASSLAIEDRGKVTITTKDDPYINVDGMSPPQGEVF 60
QY 61 VTDEGAETDLDLGHYRFLKNTTKKNNFTTGQYVEGLNERRGDTGATGTYIPIITD 120

DB 61 VTEDGAEFDLDLGHIERFLKTTMTKNNFTTGQYVEQLRNERKGVLTGATVQYIPHTD 120
QY 121 EIKRRVYSASGKQVAL:EVGGTVGDIESLPFLETIRQMGVELGRDALFIHLTVYIK 180
DB 121 EIKRRVYSASGKQVAL:EVGGTVGDIESLPFLETIRQMGVELGRDALFIHLTVYIK 180
QY 181 SAGELKTPQTHSVYELTTIGIOPDILICRSEQPIPSERKIALFTNVAEKAVISAIDA 240
DB 181 SAGELKTPQTHSVYELTTIGIOPDILICRSEQPIPSERKIALFTNVAEKAVISAIDA 240
QY 241 DTIYRIPLLRBOGLDLDLVNDQLRDLVPAADLSAMEKVVDTLPTDVSIAIVGYVDH 300
DB 241 DTIYRIPLLRBOGLDLDLVNDQLRDLVPAADLSAMEKVVDTLPTDVSIAIVGYVDH 300
QY 301 TDVAYSALBEALIHAGIHRHKVQISYIDSETEIEAGTAKLKNVDALLVPGFGRGVEGK 360
DB 301 TDVAYSALBEALIHAGIHRHKVQISYIDSETEIEAGTAKLKNVDALLVPGFGRGVEGK 360
QY 361 ISTVFAFENKIPYLGICLGMQSAVIEPARVVGLEGAHSTFPLPKSPHPIGLITEMD 420
DB 361 ISTVFAFENKIPYLGICLGMQSAVIEPARVVGLEGAHSTFPLPKSPHPIGLITEMD 420
QY 421 EAGELVTDESDLDGSTRRLGAOKCRKLAADSLAFOLYQKDVITERHHRHREYFNNQYIKOL 480
DB 421 EAGELVTDESDLDGSTRRLGAOKCRKLAADSLAFOLYQKDVITERHHRHREYFNNQYIKOL 480
QY 481 EAAGKPSGKSLDGLRVLVEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540
DB 481 EAAGKPSGKSLDGLRVLVEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540
QY 541 OGTA 544
DB 541 OGTA 544

RESULT 2
US-09-934-868-70
Sequence 70, Application US/09934868
Patent No. US20020137190A1
GENERAL INFORMATION:
APPLICANT: Koffas, Mattheos
APPLICANT: Odum, James M.
APPLICANT: Schenzle, Andreas J.
TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: CIL596 US NA
CURRENT APPLICATION NUMBER: US/09/934, 868
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft Office 97
SEQ ID NO 70
LENGTH: 544
TYPE: PRT
ORGANISM: Methylomonas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by ORF6 - PYRG
US-09-934-868-70

Query Match 100.0%; Score 2788; DB 10; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.5e-254;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKPFIITGVVSSIGKGIASSLAAILIEDRGLKVTITKDDPYTNVPGTMSPPQGEVF 60
DB 1 MTKPFIITGVVSSIGKGIASSLAAILIEDRGLKVTITKDDPYTNVPGTMSPPQGEVF 60
QY 61 VTEDGAEFDLDLGHIERFLKTTMTKNNFTTGQYVEQLRNERKGVLTGATVQYIPHTD 120
DB 61 VTEDGAEFDLDLGHIERFLKTTMTKNNFTTGQYVEQLRNERKGVLTGATVQYIPHTD 120
QY 121 EIKRRVYSASGKQVAL:EVGGTVGDIESLPFLETIRQMGVELGRDALFIHLTVYIK 180

DB 121 EIKRRVYSASGKQVAL:EVGGTVGDIESLPFLETIRQMGVELGRDALFIHLTVYIK 180
QY 181 SAGELKTPQTHSVYELTTIGIOPDILICRSEQPIPSERKIALFTNVAEKAVISAIDA 240
DB 181 SAGELKTPQTHSVYELTTIGIOPDILICRSEQPIPSERKIALFTNVAEKAVISAIDA 240
QY 241 DTIYRIPLLRBOGLDLDLVNDQLRDLVPAADLSAMEKVVDTLPTDVSIAIVGYVDH 300
DB 241 DTIYRIPLLRBOGLDLDLVNDQLRDLVPAADLSAMEKVVDTLPTDVSIAIVGYVDH 300
QY 301 TDVAYSALBEALIHAGIHRHKVQISYIDSETEIEAGTAKLKNVDALLVPGFGRGVEGK 360
DB 301 TDVAYSALBEALIHAGIHRHKVQISYIDSETEIEAGTAKLKNVDALLVPGFGRGVEGK 360
QY 361 ISTVFAFENKIPYLGICLGMQSAVIEPARVVGLEGAHSTFPLPKSPHPIGLITEMD 420
DB 361 ISTVFAFENKIPYLGICLGMQSAVIEPARVVGLEGAHSTFPLPKSPHPIGLITEMD 420
QY 421 EAGELVTDESDLDGSTRRLGAOKCRKLAADSLAFOLYQKDVITERHHRHREYFNNQYIKOL 480
DB 421 EAGELVTDESDLDGSTRRLGAOKCRKLAADSLAFOLYQKDVITERHHRHREYFNNQYIKOL 480
QY 481 EAAGKPSGKSLDGLRVLVEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540
DB 481 EAAGKPSGKSLDGLRVLVEIIELEPHWFLACQFHPFTSTPRNGHALPSGFVEAAAKHT 540
QY 541 OGTA 544
DB 541 OGTA 544

RESULT 3
US-09-941-947a-16
Sequence 16, Application US/09941947A
Publication No. US20030003528A1
GENERAL INFORMATION:
APPLICANT: Brzostowicz, Patricia C.
APPLICANT: Cheng, Qiong
APPLICANT: DiCosimo, Deana J.
APPLICANT: Koffas, Mattheos
APPLICANT: Miller, Edward S. Jr.
APPLICANT: Odum, J. Martin
APPLICANT: Picataggio, Steve
TITLE OF INVENTION: CARBON DIOXIDE PRODUCTION FROM A SINGLE CARBON SOURCE
FILE REFERENCE: CIL903 US NA
CURRENT APPLICATION NUMBER: US/09/941, 947A
CURRENT FILING DATE: 2001-09-01
PRIOR APPLICATION NUMBER: 60/229, 907
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 858
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 544
TYPE: PRT
ORGANISM: Methylomonas 16a
US-09-941-947a-16

Query Match 100.0%; Score 2788; DB 11; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.5e-254;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKPFIITGVVSSIGKGIASSLAAILIEDRGLKVTITKDDPYTNVPGTMSPPQGEVF 60
DB 1 MTKPFIITGVVSSIGKGIASSLAAILIEDRGLKVTITKDDPYTNVPGTMSPPQGEVF 60
QY 61 VTEDGAEFDLDLGHIERFLKTTMTKNNFTTGQYVEQLRNERKGVLTGATVQYIPHTD 120
DB 61 VTEDGAEFDLDLGHIERFLKTTMTKNNFTTGQYVEQLRNERKGVLTGATVQYIPHTD 120

```

QY 121 EIKRRVYESABGKQVALLIEVGSTGNDISLPLETIRMGVYELGDRALFTHLTVPIK 180
DB 121 EIKRRVYESABGKQVALLIEVGSTGNDISLPLETIRMGVYELGDRALFTHLTVPIK 180
QY 181 SAGELTKPTQHSVLEKLTIGIQPDLICRSEQPIASERRKIALFTVNAKAVISAIDA 240
DB 181 SAGELTKPTQHSVLEKLTIGIQPDLICRSEQPIASERRKIALFTVNAKAVISAIDA 240
QY 241 DITIRIPILLRQGLDDELVDQLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
DB 241 DITIRIPILLRQGLDDELVDQLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
QY 301 TDAYSLSNEALIHAGIHTRHVKQIISYIDSETTEAGTKAKLVNVALIVPGGFGRRGEVK 360
DB 301 TDAYSLSNEALIHAGIHTRHVKQIISYIDSETTEAGTKAKLVNVALIVPGGFGRRGEVK 360
QY 361 ISTVAPARENKIPYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSEHPVIGLITEMD 420
DB 361 ISTVAPARENKIPYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSEHPVIGLITEMD 420
QY 421 EAGELVTRDESDLGTRRLGQKCRKLKADSLAPOLYOKDVIYTERHRRHRYEENNOYKOL 480
DB 421 EAGELVTRDESDLGTRRLGQKCRKLKADSLAPOLYOKDVIYTERHRRHRYEENNOYKOL 480
QY 481 EAAQKPFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 540
DB 481 EAAQKPFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 540
QY 541 OGTA 544
DB 541 OGTA 544

```

RESULT 4

```

US-09-815-242-11933
; Sequence 11933, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11933
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11933

```

Query Match 72.0%; Score 2008; DB 9; Length 542;
 Best Local Similarity 71.2%; Pred. No. 8.5e-181;
 Matches 383; Conservative 64; Mismatches 91; Indels 0; Gaps 0;

```

QY 1 MKRFPIITGVVSSIGKGIKAASLAAIEDRGAKYITKLDPIYINVDPGTMSPFQGVYF 60
DB 1 MKRFPIITGVVSSIGKGIKAASLAAIEDRGAKYITKLDPIYINVDPGTMSPFQGVYF 60
QY 61 VTDEGAEITDLDGHTERFLKTTMTKKNFTTQGVYEOYLBNRKGDTGATVQVIPHITD 120
DB 61 VTDEGAEITDLDGHTERFLKTTMTKKNFTTQGVYEOYLBNRKGDTGATVQVIPHITD 120
QY 121 EIKRRVYESABGKQVALLIEVGSTGNDISLPLETIRMGVYELGDRALFTHLTVPIK 180
DB 121 EIKRRVYESABGKQVALLIEVGSTGNDISLPLETIRMGVYELGDRALFTHLTVPIK 180
QY 181 SAGELTKPTQHSVLEKLTIGIQPDLICRSEQPIASERRKIALFTVNAKAVISAIDA 240
DB 181 SAGELTKPTQHSVLEKLTIGIQPDLICRSEQPIASERRKIALFTVNAKAVISAIDA 240
QY 241 DITIRIPILLRQGLDDELVDQLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
DB 241 DITIRIPILLRQGLDDELVDQLRLDVPADLSANEKVYDGLTHEPTDEVSIAIVGYVDH 300
QY 301 TDAYSLSNEALIHAGIHTRHVKQIISYIDSETTEAGTKAKLVNVALIVPGGFGRRGEVK 360
DB 301 TDAYSLSNEALIHAGIHTRHVKQIISYIDSETTEAGTKAKLVNVALIVPGGFGRRGEVK 360
QY 361 ISTVAPARENKIPYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSEHPVIGLITEMD 420
DB 361 ISTVAPARENKIPYIGICLGMQSAVIEPARNVVGLGASHTEFLPKSEHPVIGLITEMD 420
QY 421 EAGELVTRDESDLGTRRLGQKCRKLKADSLAPOLYOKDVIYTERHRRHRYEENNOYKOL 480
DB 421 EAGELVTRDESDLGTRRLGQKCRKLKADSLAPOLYOKDVIYTERHRRHRYEENNOYKOL 480
QY 481 EAAQKPFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 538
DB 481 EAAQKPFSGKSLDGLVLEIIELEPHPMFLACQFHEFTSTPRNGHALFSGFVBAARHKT 538

```

RESULT 5

```

US-09-815-242-10275
; Sequence 10275, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```

NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10275
LENGTH: 545
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10275

Query Match 68.7%; Score 1916; DB 9; Length 545;
Best Local Similarity 67.7%; Pred. No. 4.3e-172;
Matches 363; Conservative 71; Mismatches 102; Indels 0; Gaps 0;

QY 2 TKFTFTGGVSSISGKGLAASLALIEDGKATTKLPDPYINVPDGTSPFOHGEVY 61
DB 3 TNYLFTGGVSSISGKGLAASLALIEDGKATTKLPDPYINVPDGTSPFOHGEVY 62
QY 62 TEDGAEFDLDGHTERFLTKMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPHTDE 121
DB 63 TEDGAEFDLDGHTERFLTKMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPHTDE 122
QY 122 IKRKYAESBEGKVALLEVGCTVDDISLPFLTRKQVLEGRDALPHTLVPIYKS 181
DB 123 IKRKYAESBEGKVALLEVGCTVDDISLPFLTRKQVLEGRDALPHTLVPIYKS 182
QY 182 AGEKTKPTQHSVKELERTIGIQPILICRSFOPIPASERRKIALFTVAEKAVISAIDAD 241
DB 183 AGEKTKPTQHSVKELERTIGIQPILICRSFOPIPASERRKIALFTVAEKAVISAIDAD 242
QY 242 TTYRPLLRQGIQDLDVQDLRLDVPADISAMEKYVDGTHPTDEVSIALYKTYDHT 301
DB 243 SIYKIPGLKSGQHDVYICKFSLNCEPANESEVQYIPSEANVSSEVITIGWGYKTELP 302
QY 302 DAVSYNEALFHAGIHTPHKQVSIYDSETEAGTKLKNVDAIIVPGFGSGRGEVGI 361
DB 303 DAVSYNEALFHAGIHTPHKQVSIYDSETEAGTKLKNVDAIIVPGFGSGRGEVGI 362
QY 362 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPIYIGLITTEMD 421
DB 363 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPIYIGLITTEMD 422
QY 422 AGEVLTREDSDIGTMTLGAOKRLKADSLAPQYOKDVYTERHRRYEFNNQYKOLE 481
DB 423 AGEVLTREDSDIGTMTLGAOKRLKADSLAPQYOKDVYTERHRRYEFNNQYKOLE 482
QY 482 AAGMFGSKSLDGLVEIIELEPHWPLACQHPHPEFTSTPRNGHALPSGVEAAK 537
DB 483 AAGMFGSKSLDGLVEIIELEPHWPLACQHPHPEFTSTPRNGHALPSGVEAAK 538

RESULT 6
US-09-815-242-11175
Sequence 11175, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11175
LENGTH: 545
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11175

Query Match 68.5%; Score 1908.5; DB 9; Length 545;
Best Local Similarity 67.5%; Pred. No. 2.2e-171;
Matches 364; Conservative 71; Mismatches 103; Indels 1; Gaps 1;

QY 2 TKFTFTGGVSSISGKGLAASLALIEDGKATTKLPDPYINVPDGTSPFOHGEVY 61
DB 3 TNYLFTGGVSSISGKGLAASLALIEDGKATTKLPDPYINVPDGTSPFOHGEVY 62
QY 62 TEDGAEFDLDGHTERFLTKMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPHTDE 121
DB 63 TEDGAEFDLDGHTERFLTKMTKNNFTTGOVVEQVLRNERKQDYAGATVOYIPHTDE 122
QY 122 IKRKYAESBEGKVALLEVGCTVDDISLPFLTRKQVLEGRDALPHTLVPIYKS 181
DB 123 IKRKYAESBEGKVALLEVGCTVDDISLPFLTRKQVLEGRDALPHTLVPIYKS 182
QY 182 AGEKTKPTQHSVKELERTIGIQPILICRSFOPIPASERRKIALFTVAEKAVISAIDAD 241
DB 183 AGEKTKPTQHSVKELERTIGIQPILICRSFOPIPASERRKIALFTVAEKAVISAIDAD 242
QY 242 TTYRPLLRQGIQDLDVQDLRLDVPADISAMEKYVDGTHPTDEVSIALYKTYDHT 301
DB 243 SIYKIPGLKSGQHDVYICKFSLNCEPANESEVQYIPSEANVSSEVITIGWGYKTELP 302
QY 302 DAVSYNEALFHAGIHTPHKQVSIYDSETEAGTKLKNVDAIIVPGFGSGRGEVGI 361
DB 303 DAVSYNEALFHAGIHTPHKQVSIYDSETEAGTKLKNVDAIIVPGFGSGRGEVGI 362
QY 362 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPIYIGLITTEMD 421
DB 363 STVPARENKIPYLGICLQWQSAVIEPARNVGLGASHTEFLPKSHPIYIGLITTEMD 422
QY 422 AGEVLTREDSDIGTMTLGAOKRLKADSLAPQYOKDVYTERHRRYEFNNQYKOLE 481
DB 423 AGEVLTREDSDIGTMTLGAOKRLKADSLAPQYOKDVYTERHRRYEFNNQYKOLE 482
QY 482 AAGMFGSKSLDGLVEIIELEPHWPLACQHPHPEFTSTPRNGHALPSGVEAA-ARHK 539
DB 483 AAGMFGSKSLDGLVEIIELEPHWPLACQHPHPEFTSTPRNGHALPSGVEAA-ARHK 541

RESULT 7
US-09-815-242-13865
Sequence 13865, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A


```
Db      471 8Q0LEANGV1SGTSPDRLVEMVEIPIINDFFIACQHPBELSRPNRPHPIFKSFIASL 530
Qy      537 KEK 539
Db      531 KYQ 533

RESULT 9
US-09-815-242-12978
/ Sequence 12978, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12978
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-09-815-242-12978

Query Match      56.5%; Score 1576; DB 9; Length 536;
Best Local Similarity 55.6%; Pred. No. 6e-140;
Matches 302; Conservative 93; Mismatches 134; Indels 14; Gaps 4;

Qy      1 MTKPIFTGTVSSIGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
Db      1 MTKPIFTGTVSSIGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
Qy      61 VTDDGAETDLDLGHYERELDKTTTKKONFTTGQYBEVLRNKRKGDYLGATVQYIPIHTD 120
Db      61 VTDDGAETDLDLGHYERELDKTTTKKONFTTGQYBEVLRNKRKGDYLGATVQYIPIHTD 120
Qy      121 EIKRVRYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHLTLVPY 178
Db      121 EIKRRLLAGSSTNADVITRIGTGTGIBSLPFLIRAIROQRSDIGRENVVYHCITLPPY 180
Qy      179 IKSAGELTKPTQSHVKEKRLPTIGTOPDLIRKSGOPIPASRRKIALFTVNAEKVYSAI 238
Db      181 IKAAGEKTKPTQSHVKEKRLPTIGTOPDLIRKSGOPIPASRRKIALFTVNAEKVYSAI 240
Qy      239 DADITIRPLILEEGGLDVLVDQIRLDPV--AADISAMEKVVDELTPTEVSIAIVGKI 297
Db      241 DADISYELPILQSGQNMDDIVIKRLQINAKYETQLDWEKQLDDIVNNLDKKITIGLVGKI 300
Qy      298 VDHDAVYSALNALIHAGIHRKRVQIYSIINSETIEAGRAK--LKNVDALIVPGRGSEK 356
```

```
Db      301 VSLQDAIYLSVBSLKHAAIPPAKQIDIRIMISSSVTDENAEYLAADVIGLIVPGGFRA 360
Qy      357 VEGKISTVPARENKIPYLGIQLGQSAVIEPARNVGLGASHTEPLPKSEHPIGLIT 416
Db      361 SEKGISALIKYARENVPFGICLQWOLATVYFSHNVIGLEGAHSAELDPATPYDILLLP 420
Qy      417 EEMDAGELVTRDESDIGTMRPLGAQCRKADSLAFOLYQXVITEERHRYEFPNNQY 476
Db      421 ECKD-----LEDIGTLRLGLTFCSTIKEGTLADQVYGAKEIERHRHRIEFPNDY 470
Qy      477 EKQLEPAQKESGKSLDRIVEIIELEPHWFLACQHPPEFTSTPRNGHALFSGFVAAA 536
Db      471 8Q0LEANGV1SGTSPDRLVEMVEIPIINDFFIACQHPBELSRPNRPHPIFKSFIASL 530
Qy      537 KEK 539
Db      531 KYQ 533

RESULT 10
US-09-815-242-4984
/ Sequence 4984, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 4984
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-815-242-4984

Query Match      56.3%; Score 1569; DB 9; Length 536;
Best Local Similarity 55.4%; Pred. No. 2.8e-139;
Matches 301; Conservative 94; Mismatches 134; Indels 14; Gaps 4;

Qy      1 MTKPIFTGTVSSIGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
Db      1 MTKPIFTGTVSSIGKGIASLSAILEDRLVTTTKLDPIYINVPDGTMSPPQHGVEF 60
Qy      61 VTDDGAETDLDLGHYERELDKTTTKKONFTTGQYBEVLRNKRKGDYLGATVQYIPIHTD 120
Db      61 VTDDGAETDLDLGHYERELDKTTTKKONFTTGQYBEVLRNKRKGDYLGATVQYIPIHTD 120
Qy      121 EIKRVRYSAS--GKDVALLIEVGTVGDIBSLPFLFTIRQMGVZLGRDRLFIHLTLVPY 178
```

```

Db      121 EIKKXIMRAKMTDADVITTEVGTGVIIESLPFEALROKMGSDONVYIHTTLPY 180
Qy      179 IKSAGELKTPQHSVKEKRTIGIQPDILICSEQPIPASERRKIALFTNVAEKAVISAI 238
Db      181 LKAAGEMKTPQHSVKEKRTIGIQPDILICSEQPIPASERRKIALFTNVAEKAVISAI 240
Qy      239 DADTYRIPLLREGLDDLVVDQRLVPAADLSAMERKVDGLTHPTDEVIAIVGKYV 298
Db      241 DVETLYSIPALQANMOQIVCDHLKLPADMTMTRALEKXVNLKKTKIALVGYV 300
Qy      299 DHTDVKSLNEALIHAGIHTRKVOISYIDSETIABGTA-KLKXVDAIIVPGGEGRGV 357
Db      301 ELPDAIVSVEALKAGPDPDSIDIDWVDSOELTAENVAERIGSADGILVPGGEGRG 360
Qy      358 ECKISTVPARENKIPYIGICIGMOSAVIEFARNVVGEGASTEPFKSPHPVIGLITE 417
Db      361 ECKIEIRPARENDVPFGICIGMOMACVEFGRNVGLEDAGSAETNPDVNNITIDMAD 420
Qy      418 NMDAGELVTRDESDIGTMTLGOKRLKADSLAPOLY-QKQVITERHRHRYEFNNQY 476
Db      421 -----QENINLGGTTLGLGYPKLKKGTAAAYGNEDVQRRHRRYEFNNKY 470
Qy      477 LKOLEAGKFFSGKSLDGLVYIIELEPHWFLACQFHPFTSPRNHALFSGFEVAAA 536
Db      471 RQLFENGIVFSGVSPDRNLVIVAIPEKQFVACQFHPILISRNRRQLIKGFVGAAL 530
Qy      537 KHK 539
Db      531 ANK 533

RESULT 11
US-09-815-242-10727
/ Sequence 10727, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815.242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 10727
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-815-242-10727
Query Match          56.3%; Score 1569; DB 9; Length 536;

```

```

Best Local Similarity 55.4%; Pred. No. 2.8e-139;
Matches 301; Conservative 94; Mismatches 134; Indels 14; Gaps 4;

Qy      1 MKKFIPTGVSSTGKGLAASLAIEDRLKVTITKLPDYINVDQITMSPFQHWVF 60
Db      1 MKKFIPTGVSSTGKGLVAAISLRLKRLKRLKRTIKQFDPYINVDQITMSPFQHWVF 60
Qy      61 VTBDGAEITDIGHYERPLKTTMTKNNFTTGQVYEQVLRNBRKGDYIGATVQVPIHTTD 120
Db      61 VTBDGAEITDIGHYERFDINLANKYSVLTGKISYVLRKRKGEYIGATVQVPIHTTN 120
Qy      121 EIKRKYAESAE--GQDVALLIEVGTGSGDIESLPFETIROKGVLEGRRLFIHTLTPY 176
Db      121 EIKKXIMRAKMTDADVITTEVGTGVIIESLPFEALROKMGSDONVYIHTTLPY 180
Qy      179 IKSAGELKTPQHSVKEKRTIGIQPDILICSEQPIPASERRKIALFTNVAEKAVISAI 238
Db      181 LKAAGEMKTPQHSVKEKRTIGIQPDILICSEQPIPASERRKIALFTNVAEKAVISAI 240
Qy      239 DADTYRIPLLREGLDDLVVDQRLVPAADLSAMERKVDGLTHPTDEVIAIVGKYV 298
Db      241 DVETLYSIPALQANMOQIVCDHLKLPADMTMTRALEKXVNLKKTKIALVGYV 300
Qy      299 DHTDVKSLNEALIHAGIHTRKVOISYIDSETIABGTA-KLKXVDAIIVPGGEGRGV 357
Db      301 ELPDAIVSVEALKAGPDPDSIDIDWVDSOELTAENVAERIGSADGILVPGGEGRG 360
Qy      358 ECKISTVPARENKIPYIGICIGMOSAVIEFARNVVGEGASTEPFKSPHPVIGLITE 417
Db      361 ECKIEIRPARENDVPFGICIGMOMACVEFGRNVGLEDAGSAETNPDVNNITIDMAD 420
Qy      418 NMDAGELVTRDESDIGTMTLGOKRLKADSLAPOLY-QKQVITERHRHRYEFNNQY 476
Db      421 -----QENINLGGTTLGLGYPKLKKGTAAAYGNEDVQRRHRRYEFNNKY 470
Qy      477 LKOLEAGKFFSGKSLDGLVYIIELEPHWFLACQFHPFTSPRNHALFSGFEVAAA 536
Db      471 RQLFENGIVFSGVSPDRNLVIVAIPEKQFVACQFHPILISRNRRQLIKGFVGAAL 530
Qy      537 KHK 539
Db      531 ANK 533

RESULT 12
US-09-815-242-11376
/ Sequence 11376, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815.242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931

```

PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 13376
 LENGTH: 535
 TYPE: PR
 ORGANISM: Streptococcus pneumoniae
 US-09-815-242-13376

Query Match 56.0%; Score 1561; DB 9; Length 535;
 Best Local Similarity 54.3%; Pred. No. 1,6e-138;
 Matches 294; Conservative 108; Mismatches 125; Indels 14; Gaps 4;

QY 2 TKTFTFGVSSSGKGIASLAILEDRLAVTITKDPYINVDPTGMSPFQHGEPVT 61
 DB 3 TKYFTVGAVSSSGKGIASLAILEDRLAVTITKDPYINVDPTGMSPFQHGEPVT 62
 QY 62 TEDGAEFDLDGHERFLKTTMTKNNFTGQVEQVLRNKRKDYAGATVOYPIHTDE 121
 DB 63 TDDGAEFDLDGHERFLDINLNKYSNVTGKIYSEVLRKREGEREYLGAIVQYPIHTDA 122
 QY 122 IKRNVESA--EGQDVALLIEVGSTVGDIESTPLEFTIRQMGVLEGRDALFTHTLTVPTI 179
 DB 123 LKEXIKRAALTDSVITTEVGTVGISLPLEFLALRQKRAVDGADNVMYIHTLITPL 182
 QY 180 KSADELKPTQSHSVKELRTIGIOPDLICRSQPIPBASERRKIALFTVAEKAVIDA 239
 DB 183 KKAQBMKTKPIQHSVKELRGIGIOPNLVIRTEBPAGGIKNLAFCOVAPEAVIESID 242
 QY 240 ADLTIRIPLIREGDDLVVDQRLDVPADLSAMEKVVDGLTPTDEVSIALVSKYVD 299
 DB 243 VEHLYQIPLNIOQGMQIVCDHLKIDAPALDWTESAMWDKVMNLKQVKSILVGYVE 302
 QY 300 HTDVKSLNEALIHAGIHTHRKVOISYIDSETIEAGTAK-LKNVALIIVPGGFERGVE 358
 DB 303 LQDAIVISVEALKSGVYNDVEKIMVANDVTAENVALLSDAGIIVPGGFERGVE 362
 QY 359 GKISTVPARENKIPLYGICLQMSAVIEFARNVVGEGASTFELPKSPHPVIGLITTEM 418
 DB 363 GKIOAIRARENDVPMGLGVCLQMLTCEIPARHYLGIEGANSALPETYPTIIDMRQ 422
 QY 419 MDEAGELVTDESDLGIMRLGAQKCELRKADSLAPQY-QKDVITEBRRHREPNQVL 477
 DB 423 ID-----IEDWGCTRLGLYPSKLRGSAALAAVNMQEVQRRHRYEFPNNAFR 472
 QY 478 KQLEBAQMKESGKSLDRLVEIIELEPFWFLACQFHPBFTSTPRGHALFSGFVZAAK 537
 DB 473 EQFLAAGFVSGVSPDRILVEIYIENKRFVACQYHPELSSRPNRPESLYTAFVTAIVE 532
 QY 538 H 538
 DB 533 N 533

RESULT 13

US-09-815-242-11320
 Sequence 11320, Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
 APPLICANT: Olsen, Karl L.
 APPLICANT: Zvekind, Judith W.
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 11320
 LENGTH: 538
 TYPE: PR
 ORGANISM: Helicobacter pylori
 US-09-815-242-11320

Query Match 51.8%; Score 1444.5; DB 9; Length 538;
 Best Local Similarity 50.8%; Pred. No. 1.7e-127;
 Matches 271; Conservative 112; Mismatches 147; Indels 3; Gaps 2;

QY 3 KFIETFGVSSSGKGIASLAILEDRLAVTITKDPYINVDPTGMSPFQHGEPVT 62
 DB 5 KFIETFGVSSSGKGISSSIATLQHCNVOVISIKIDPYINIDPTGMSPFQHGEPVT 64
 QY 63 EDGAEFDLDGHERFLKTTMTKNNFTGQVEQVLRNKRKDYAGATVOYPIHTDE 122
 DB 65 SDGAEFDLDGHERFLNRLTRNNFTGQISSVLEMRKSGTIGKIIQYPIHTDE 124
 QY 123 KRAYASAGQDVALLIEVGSTVGDIESTPLEFTIRQMGVLEGRDALFTHTLTVPIKSA 182
 DB 125 KRAIKSAAGDPLIVVGSTVGDMSEMFLEALRQKLEMEKVINAVTLPIYIQT 184
 QY 183 GELTKPTQSHSVKELRTIGIOPDLICRSQPIPBASERRKIALFTVAEKAVIDA 242
 DB 185 NELTKPTQSHSVKELRLAVTPOIILASPKPLDKELKNLALSCVQDSVVAATDTS 244
 QY 243 IYRPIPLIREGDDLVVDQRLDVPADLSAMEKVVDGLTPTDEVSIALVSKYVD 302
 DB 245 IYACPILFLQSGIITPLARRPNLKLHPMAANTLVEKILAKHVKIGFVGKYLSE 304
 QY 303 AYKSLNEALIHAGIHTHRKVOISYIDSETIEAGTAKLKNVDAIIVPGGFERGVE 362
 DB 305 SYKSLIENALIHAGIHTDQVNIEMIDSENFNER-TDLEGVDAIIVPGGFERGVE 362
 QY 363 TVRPARENKIPLYGICLQMSAVIEFARNVVGEGASTFELPKSPHPVIGLITTEM 422
 DB 363 ALQPARLEKLPPLDGLIQLQMSALIVERRNALGKANSSTFNRQCEVPVYLLGDMDQ 422
 QY 423 GELVTDESDLGIMRLGAQKCELRKADSLAPQYQKDVITEBRRHREPNQVL 482
 DB 423 HQQVRYTNSPLIGSTMLGEYCEBIMNSLLEKAYKCPISKEHRRHRYELNPKRDEWEN 482
 QY 483 AGMFGSKSLDGLVEIIELEPFWFLACQFHPBFTSTPRGHALFSGFVZAAK 535
 DB 483 KGLNVGPG-SNHLIEALIEHDHFPVGVQFHPBFTSRQSPNPTIIDPFXSA 534

RESULT 14

US-10-156-761-14031
 Sequence 14031, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI

```

; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 14031
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14031

```

```

Query Match          51.4%; Score 1432.5; DB 15; Length 549;
Best Local Similarity 52.5%; Pred. No. 2,3e-126;
Matches 287; Conservative 92; Mismatches 153; Indels 15; Gaps 7;

```

```

QY 2 TKFIFITGCVSSSLCKGIAASSLAILEDRLKVTITKLDPIYINVDGTVSPFGHGEV 61
DB 8 TKHIFVTGCVSSSLCKGIAASSLAILEDRLKVTITKLDPIYINVDGTVSPFGHGEV 67
QY 62 TEDGATDLGHTERFLKTMKNNFTTGOVYEOVLNRKGDYIGATVQVPHITDE 121
DB 68 TNDGATDLGHTERFLDRDLDSANVTGOVYVIAERKRGVYAGDVQVPHITDE 127
QY 122 IKRKYVESAEK-DVALIEVGTVGDIESLPLETIRQMGVELGRDALFIHLLTVYIK 180
DB 128 IKHRRBATTEVDVYIEVGTVGDIESLPLETIRQMGVELGRDALFIHLLTVYIK 187
QY 181 SAGELKTPQHSVKEITGIGOPDLICRSEOPIPASERKIALFTNVAEKAVISAIDA 240
DB 188 PSGLKTPQHSVAAKNIQIPDAIVLCREVPALIKRKISIMCDDVBAVACPDA 247
QY 241 DTIYRPLLRBOGLDVLVDQLRDLVPAADLSAMEKVVDGLTPTDEVSIAIVGKYVDH 300
DB 248 RSIYDIPKTVGEGDGLAVYAKLDPFVDVDTWTDLDLDRVKEPHEITLALVGTIDL 307
QY 301 TDAYKSLNBLIHAGIHTRHVQISYIDSEITIE--ABGTAKLKNVDAIIVPGGGRGV 358
DB 308 PDAYLSVTEALPAGGFARARVXIKWVTSDDCKTPAGAGQLGVDALICIPGFGRGV 367
QY 359 GKISVFRARENKTPYIGICGMSAVIEFARNVGLGASHTFELPKSFHPIGILTEM 418
DB 368 GKVAGIQYAREHRIPLLGLCLGQCIYIEARNLADIIPANSTEFDSATGHPVISTMAAQ 427
QY 419 MD-EAGSLVTRDESDSLGTMRLGAOKCRLKADSLAFQLYQ-KDVIERRHRYERNNOY 476
DB 428 LDIVAGS-----GDMGTMRLGMYPALKLAESIVABVDKGYEERHRYERNNOY 480
QY 477 LKQLE-AAAGKFSKSLDGLVETIIEPE--HMFPLACOPHPEFTSPRNGALLFGVYE 533
DB 481 RAELBKRAGLDFSGTSDGKLVYEVYEPREHVPYLVATQAHPELRSPTEPHPLFAGLVK 540
QY 534 AAAXHKT 540
DB 541 AAVERTK 547

```

```

RESULT 15
US-09-738-626-5066
; Sequence 5066, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

```

```

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 5066
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5066

```

```

Query Match          49.2%; Score 1372; DB 10; Length 554;
Best Local Similarity 51.9%; Pred. No. 1,3e-120;
Matches 283; Conservative 78; Mismatches 166; Indels 18; Gaps 6;

```

```

QY 2 TKFIFITGCVSSSLCKGIAASSLAILEDRLKVTITKLDPIYINVDGTVSPFGHGEV 61
DB 10 TKHIFVTGCVSSSLCKGIAASSLAILEDRLKVTITKLDPIYINVDGTVSPFGHGEV 69
QY 62 TEDGATDLGHTERFLKTMKNNFTTGOVYEOVLNRKGDYIGATVQVPHITDE 121
DB 70 TNDGATDLGHTERFLDRDLDSANVTGOVYVIAERKRGVYAGDVQVPHITDE 129
QY 122 IKRKYVESAE-----GNDVALIEVGTVGDIESLPLETIRQMGVELGRDALFIHLL 175
DB 130 IKARLSNGEPDAGNADVAVISEVGTVGDIESLPLETIRQMGVELGRDALFIHLL 189
QY 176 VPIYSAGELKTPQHSVKEITGIGOPDLICRSEOPIPASERKIALFTNVAEKAVI 235
DB 190 VPIYATSGELKTPQHSVKEITGIGOPDLICRSEOPIPASERKIALFTNVAEKAVI 249
QY 236 SAIDDTIYRPLLRBOGLDVLVDQLRDLVPAADLSAMEKVVDGLTPTDEVSIAIVG 295
DB 250 SCPDSSSTIYNPVLYREHDTPTIIRLGLPFRVDSTWHDLLERVNNRPHHETVIG 309
QY 296 KYVDHTDAYKSLNBLIHAGIHTRHVQISYIDSEITIEABGTA--KLKNVDAIIVPGG 353
DB 310 KYIDLPDAVLSVVAVAAGVANNTRTNKIKITSDDCETPSGAMKOLSGLDALIVPGG 369
QY 364 ERGVGKISTVFRARENKIPYIGICGMSAVIEFARNVGLGASHTFELPKSFHPIVIG 413
DB 370 IRGIBKGIATTPREKIPILGICLGLQCTVIEAARQ-AGLEBASSTEDPDAITQVIA 428
QY 414 LITMMDAGELVTRDESDSLGTMRLGAOKCRLKADSLAFQLYQKDVITERRHRYERN 473
DB 429 TMEB-----OKXAVSGADIGTMRLGAVPATLEBSGLAELVYGFVESEHRYERN 482
QY 474 NOYLKOL-EAAGKFSKSLDGLVETIIEPE--HMFPLACOPHPEFTSPRNGALLFG 530
DB 483 NAYRAQIAESDVLFSGTSPDGLHVEVYEVYEPREHVPYLVATQAHPELRSPTEPHPL 542
QY 531 FVEAA 535
DB 543 LVKTA 547

```

```

Search completed: January 29, 2004, 16:21:18
Job time : 38.1346 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using SW model

Run on: January 29, 2004, 15:54:19 ; Search time 15.7685 Seconds
(without alignments)
3044.503 Million cell updates/sec

Title: US-09-941-947a-10

Perfect score: 1199
Sequence: 1 MNPTICQWAVVPAAGVGRKM.....IKITRPEDLALAQFYEQQA 231

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1199	100.0	231 10 US-09-934-903-6	Sequence 6, Appl1
2	1199	100.0	231 10 US-09-934-868-68	Sequence 68, Appl1
3	1199	100.0	231 11 US-09-941-947a-10	Sequence 10, Appl1
4	555.5	46.3	234 9 US-09-815-242-11931	Sequence 11931, A
5	530.5	44.2	225 9 US-09-815-242-11076	Sequence 11076, A
6	529.5	44.2	254 9 US-09-815-242-11722	Sequence 11722, A
7	520.5	43.4	236 9 US-09-815-242-10273	Sequence 10273, A
8	520.5	43.4	236 9 US-09-792-251-11	Sequence 11, Appl1
9	499.5	41.7	236 9 US-09-815-242-13867	Sequence 13867, A
10	297	24.8	232 10 US-09-792-251-8	Sequence 8, Appl1
11	283.5	23.6	302 10 US-09-935-943-8	Sequence 8, Appl1
12	255.5	21.3	232 12 US-10-128-713A-6	Sequence 6, Appl1
13	245.5	20.5	250 15 US-10-156-761-11503	Sequence 11503, A
14	240.5	20.1	218 12 US-10-289-762-621	Sequence 621, App
15	223.5	18.6	256 10 US-09-738-626-6422	Sequence 6422, Ap

5	217.5	18.1	238 9 US-09-815-242-12471	Sequence 12471, A
7	214.5	17.9	227 9 US-09-815-242-5711	Sequence 5711, Ap
3	214.5	17.9	238 9 US-09-876-181-2	Sequence 2, Appl1
9	209	17.4	235 9 US-09-815-242-13256	Sequence 13256, A
0	209	17.4	235 10 US-09-792-251-2	Sequence 2, Appl1
1	198.5	16.6	238 9 US-09-876-182-2	Sequence 2, Appl1
2	193.5	16.1	236 9 US-09-815-242-5707	Sequence 5707, Ap
3	193.5	16.1	242 9 US-09-815-242-12467	Sequence 12467, Ap
4	126	10.5	125 9 US-09-867-550-658	Sequence 658, App
5	105.5	8.8	291 12 US-10-369-493-21638	Sequence 21638, A
6	98.5	8.2	761 9 US-09-815-242-10937	Sequence 10937, A
7	98.5	8.2	483 10 US-09-738-626-5017	Sequence 5017, Ap
13	98	8.2	395 12 US-10-369-493-18528	Sequence 18528, A
9	94	7.8	327 12 US-10-369-493-20306	Sequence 20306, A
0	93	7.8	297 12 US-10-369-493-17444	Sequence 17444, A
1	93	7.8	297 12 US-10-312-273-245	Sequence 245, App
2	93	7.8	357 12 US-10-289-762-487	Sequence 487, App
3	91	7.6	419 12 US-10-369-493-21686	Sequence 21686, A
4	90.5	7.5	285 12 US-10-369-493-1153	Sequence 1153, Ap
5	90	7.5	456 9 US-09-815-242-10400	Sequence 10400, A
6	89.5	7.5	381 16 US-10-080-170-264	Sequence 264, App
7	89.5	7.5	385 12 US-10-369-493-21535	Sequence 21535, A
8	89	7.4	154 9 US-09-939-960-528	Sequence 528, App
9	89	7.4	704 15 US-10-128-714-8238	Sequence 8238, Ap
0	89	7.4	1317 12 US-10-369-493-13693	Sequence 13693, A
1	88.5	7.4	361 12 US-10-369-493-23443	Sequence 23443, A
2	88.5	7.4	485 10 US-09-738-626-4551	Sequence 4551, Ap
3	88	7.3	394 12 US-10-369-493-11824	Sequence 11824, A
4	87.5	7.3	443 15 US-10-156-761-14528	Sequence 14528, A
5	87.5	7.3	831 15 US-10-156-761-14505	Sequence 14505, A

ALIGNMENTS

REST T 1	US-0-934-903-6	Sequence 6, Application US/09934903
1	US-0-934-903-6	Sequence 6, Application US/09934903
2	US-0-934-903-6	Sequence 6, Application US/09934903
3	US-0-934-903-6	Sequence 6, Application US/09934903
4	US-0-934-903-6	Sequence 6, Application US/09934903
5	US-0-934-903-6	Sequence 6, Application US/09934903
6	US-0-934-903-6	Sequence 6, Application US/09934903
7	US-0-934-903-6	Sequence 6, Application US/09934903
8	US-0-934-903-6	Sequence 6, Application US/09934903
9	US-0-934-903-6	Sequence 6, Application US/09934903
10	US-0-934-903-6	Sequence 6, Application US/09934903
11	US-0-934-903-6	Sequence 6, Application US/09934903
12	US-0-934-903-6	Sequence 6, Application US/09934903
13	US-0-934-903-6	Sequence 6, Application US/09934903
14	US-0-934-903-6	Sequence 6, Application US/09934903
15	US-0-934-903-6	Sequence 6, Application US/09934903
16	US-0-934-903-6	Sequence 6, Application US/09934903
17	US-0-934-903-6	Sequence 6, Application US/09934903
18	US-0-934-903-6	Sequence 6, Application US/09934903
19	US-0-934-903-6	Sequence 6, Application US/09934903
20	US-0-934-903-6	Sequence 6, Application US/09934903
21	US-0-934-903-6	Sequence 6, Application US/09934903
22	US-0-934-903-6	Sequence 6, Application US/09934903
23	US-0-934-903-6	Sequence 6, Application US/09934903
24	US-0-934-903-6	Sequence 6, Application US/09934903
25	US-0-934-903-6	Sequence 6, Application US/09934903
26	US-0-934-903-6	Sequence 6, Application US/09934903
27	US-0-934-903-6	Sequence 6, Application US/09934903
28	US-0-934-903-6	Sequence 6, Application US/09934903
29	US-0-934-903-6	Sequence 6, Application US/09934903
30	US-0-934-903-6	Sequence 6, Application US/09934903
31	US-0-934-903-6	Sequence 6, Application US/09934903
32	US-0-934-903-6	Sequence 6, Application US/09934903
33	US-0-934-903-6	Sequence 6, Application US/09934903
34	US-0-934-903-6	Sequence 6, Application US/09934903
35	US-0-934-903-6	Sequence 6, Application US/09934903
36	US-0-934-903-6	Sequence 6, Application US/09934903
37	US-0-934-903-6	Sequence 6, Application US/09934903
38	US-0-934-903-6	Sequence 6, Application US/09934903
39	US-0-934-903-6	Sequence 6, Application US/09934903
40	US-0-934-903-6	Sequence 6, Application US/09934903
41	US-0-934-903-6	Sequence 6, Application US/09934903
42	US-0-934-903-6	Sequence 6, Application US/09934903
43	US-0-934-903-6	Sequence 6, Application US/09934903
44	US-0-934-903-6	Sequence 6, Application US/09934903
45	US-0-934-903-6	Sequence 6, Application US/09934903
46	US-0-934-903-6	Sequence 6, Application US/09934903
47	US-0-934-903-6	Sequence 6, Application US/09934903
48	US-0-934-903-6	Sequence 6, Application US/09934903
49	US-0-934-903-6	Sequence 6, Application US/09934903
50	US-0-934-903-6	Sequence 6, Application US/09934903
51	US-0-934-903-6	Sequence 6, Application US/09934903
52	US-0-934-903-6	Sequence 6, Application US/09934903
53	US-0-934-903-6	Sequence 6, Application US/09934903
54	US-0-934-903-6	Sequence 6, Application US/09934903
55	US-0-934-903-6	Sequence 6, Application US/09934903
56	US-0-934-903-6	Sequence 6, Application US/09934903
57	US-0-934-903-6	Sequence 6, Application US/09934903
58	US-0-934-903-6	Sequence 6, Application US/09934903
59	US-0-934-903-6	Sequence 6, Application US/09934903
60	US-0-934-903-6	Sequence 6, Application US/09934903
61	US-0-934-903-6	Sequence 6, Application US/09934903
62	US-0-934-903-6	Sequence 6, Application US/09934903
63	US-0-934-903-6	Sequence 6, Application US/09934903
64	US-0-934-903-6	Sequence 6, Application US/09934903
65	US-0-934-903-6	Sequence 6, Application US/09934903
66	US-0-934-903-6	Sequence 6, Application US/09934903
67	US-0-934-903-6	Sequence 6, Application US/09934903
68	US-0-934-903-6	Sequence 6, Application US/09934903
69	US-0-934-903-6	Sequence 6, Application US/09934903
70	US-0-934-903-6	Sequence 6, Application US/09934903
71	US-0-934-903-6	Sequence 6, Application US/09934903
72	US-0-934-903-6	Sequence 6, Application US/09934903
73	US-0-934-903-6	Sequence 6, Application US/09934903
74	US-0-934-903-6	Sequence 6, Application US/09934903
75	US-0-934-903-6	Sequence 6, Application US/09934903
76	US-0-934-903-6	Sequence 6, Application US/09934903
77	US-0-934-903-6	Sequence 6, Application US/09934903
78	US-0-934-903-6	Sequence 6, Application US/09934903
79	US-0-934-903-6	Sequence 6, Application US/09934903
80	US-0-934-903-6	Sequence 6, Application US/09934903
81	US-0-934-903-6	Sequence 6, Application US/09934903
82	US-0-934-903-6	Sequence 6, Application US/09934903
83	US-0-934-903-6	Sequence 6, Application US/09934903
84	US-0-934-903-6	Sequence 6, Application US/09934903
85	US-0-934-903-6	Sequence 6, Application US/09934903
86	US-0-934-903-6	Sequence 6, Application US/09934903
87	US-0-934-903-6	Sequence 6, Application US/09934903
88	US-0-934-903-6	Sequence 6, Application US/09934903
89	US-0-934-903-6	Sequence 6, Application US/09934903
90	US-0-934-903-6	Sequence 6, Application US/09934903
91	US-0-934-903-6	Sequence 6, Application US/09934903
92	US-0-934-903-6	Sequence 6, Application US/09934903
93	US-0-934-903-6	Sequence 6, Application US/09934903
94	US-0-934-903-6	Sequence 6, Application US/09934903
95	US-0-934-903-6	Sequence 6, Application US/09934903
96	US-0-934-903-6	Sequence 6, Application US/09934903
97	US-0-934-903-6	Sequence 6, Application US/09934903
98	US-0-934-903-6	Sequence 6, Application US/09934903
99	US-0-934-903-6	Sequence 6, Application US/09934903
100	US-0-934-903-6	Sequence 6, Application US/09934903
101	US-0-934-903-6	Sequence 6, Application US/09934903
102	US-0-934-903-6	Sequence 6, Application US/09934903
103	US-0-934-903-6	Sequence 6, Application US/09934903
104	US-0-934-903-6	Sequence 6, Application US/09934903
105	US-0-934-903-6	Sequence 6, Application US/09934903
106	US-0-934-903-6	Sequence 6, Application US/09934903
107	US-0-934-903-6	Sequence 6, Application US/09934903
108	US-0-934-903-6	Sequence 6, Application US/09934903
109	US-0-934-903-6	Sequence 6, Application US/09934903
110	US-0-934-903-6	Sequence 6, Application US/09934903
111	US-0-934-903-6	Sequence 6, Application US/09934903
112	US-0-934-903-6	Sequence 6, Application US/09934903
113	US-0-934-903-6	Sequence 6, Application US/09934903
114	US-0-934-903-6	Sequence 6, Application US/09934903
115	US-0-934-903-6	Sequence 6, Application US/09934903
116	US-0-934-903-6	Sequence 6, Application US/09934903
117	US-0-934-903-6	Sequence 6, Application US/09934903
118	US-0-934-903-6	Sequence 6, Application US/09934903
119	US-0-934-903-6	Sequence 6, Application US/09934903
120	US-0-934-903-6	Sequence 6, Application US/09934903

```

Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLYHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Qy      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231

```

```

RESULT 2
US-09-934-868-68
Sequence 68: Application US/09941947A
Patent No. US20020137190A1
GENERAL INFORMATION:
APPLICANT: Koffas, Mattheos
APPLICANT: Odom, James M
APPLICANT: Schenzle, Andreas J
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: CL1936 US NA
CURRENT APPLICATION NUMBER: US/09/934, 868
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229, 858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft Office 97
SEQ ID NO: 68
LENGTH: 231
TYPE: PRT
ORGANISM: Methylomonas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by ISPD
US-09-934-868-68

```

```

Query Match      100.0%; Score 1199; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTLTLLSDFQKVAVAISVE 60
Db      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTLTLLSDFQKVAVAISVE 60
Qy      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLYHDAAPCLTGSDI 120
Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLYHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Qy      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231

```

```

RESULT 3
US-09-941-947a-10
Sequence 10: Application US/09941947A
Patent No. US20030003528A1
GENERAL INFORMATION:
APPLICANT: Bryzotowicz, Patricia C.
APPLICANT: Cheng, Qiong
APPLICANT: DiCosimo, Deana J.
APPLICANT: Koffas, Mattheos
APPLICANT: Miller, Edward S. Jr.
APPLICANT: Odom, J. Martin
APPLICANT: Picataggio, Steve
APPLICANT: Rouviere, Pierre E.
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
FILE REFERENCE: CL1903 US NA
CURRENT APPLICATION NUMBER: US/09/941, 947A

```

```

CURRENT FILING DATE: 2001-09-01
ROR APPLICATION NUMBER: 60/229, 907
ROR FILING DATE: 2000-09-01
ROR APPLICATION NUMBER: 60/229, 858
ROR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Microsoft Office 97
SEQ ID NO: 10
LENGTH: 231
TYPE: PRT
ORGANISM: Methylomonas 16a
US-9-941-947a-10

```

```

Query Match      100.0%; Score 1199; DB 11; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTLTLLSDFQKVAVAISVE 60
Db      1 MNPTIQCAVAVPAAGVGRKQADRPQYLPAGKTVIEHTLTLLSDFQKVAVAISVE 60
Qy      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLYHDAAPCLTGSDI 120
Db      61 DPWPELSIAKHPIITAPGKERADSVLSALKALEDIASNDWVLYHDAAPCLTGSDI 120
Qy      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Db      121 HLOIDTLKNDPVGGIILASSHDTLKAVDGTITATIDRKHWRAALTPOPKYGMRLDALQ 180
Qy      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231
Db      181 RTEGNPAVTDEASALIELGHKPKIVEGRPNIKITRPEDLALAQFYMEQQA 231

```

```

REF ID 4
US-9-815-242-11931
Sequence 11931: Application US/09845242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zysek, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trivick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCES: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815, 242
CURRENT FILING DATE: 2001-03-21
ROR APPLICATION NUMBER: 60/191, 078
ROR FILING DATE: 2000-03-21
ROR APPLICATION NUMBER: 60/206, 848
ROR FILING DATE: 2000-05-23
ROR APPLICATION NUMBER: 60/207, 727
ROR FILING DATE: 2000-05-26
ROR APPLICATION NUMBER: 60/242, 578
ROR FILING DATE: 2000-10-23
ROR APPLICATION NUMBER: 60/253, 625
ROR FILING DATE: 2000-11-27
ROR APPLICATION NUMBER: 60/257, 931
ROR FILING DATE: 2000-12-22
ROR APPLICATION NUMBER: 60/269, 308
ROR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: PaeSeQ for Windows Version 4.0
SEQ ID NO: 11931
LENGTH: 234
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-9-815-242-11931

```

Query Match 46.3%; Score 555.5; DB 9; Length 234;
 Best Local Similarity 54.1%; Pred. No. 5.3e-45;
 Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;

QY 8 WAWPAAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 67
 DB 10 WTVIPAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 69

QY 68 SIAGPDITTAGGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 126
 DB 70 DCAASRHYQRAAGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 129

QY 127 LKNDVGGILASSDHTLKHVDG-TITATIDRKHYVRLTPQWFKYGMALDALQ-TEG 184
 DB 130 LAEDVGGILAVPADITLKSDRDSRVSITIDRSYVVLATPQWFKYGMALDALQ-TEG 189

QY 185 NPATDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 224
 DB 190 GVAITDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 229

RESULT 5

US-09-815-242-11076
 Sequence 11076, Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815.242
 PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for windows Version 4.0
 SEQ ID NO 11076
 LENGTH: 225
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 US-09-815-242-11076

Query Match 44.2%; Score 530.5; DB 9; Length 225;
 Best Local Similarity 52.5%; Pred. No. 1.9e-46;
 Matches 117; Conservative 28; Mismatches 65; Indels 13; Gaps 6;

QY 9 AVPAAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 68
 DB 7 AVPAAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 66

QY 69 IAKHPDITTAGGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 128
 DB 67 L-DEKIDITTAGGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 117

QY 129 --NDVGGILASSDHTLKHVDG-TITATIDRKHYVRLTPQWFKYGMALDALQ-TEG 184
 DB 118 AIEDQAGILAVPADITLKSDRDSRVSITIDRSYVVLATPQWFKYGMALDALQ-TEG 177

QY 185 NPATDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 227
 DB 178 GANITDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 220

RESULT 6

US-09-815-242-11722
 Sequence 11722, Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815.242
 PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for windows Version 4.0
 SEQ ID NO 11722
 LENGTH: 254
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-815-242-11722

Query Match 44.2%; Score 529.5; DB 9; Length 254;
 Best Local Similarity 51.1%; Pred. No. 2.9e-46;
 Matches 114; Conservative 30; Mismatches 70; Indels 9; Gaps 5;

QY 9 AVPAAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 68
 DB 33 AVPAAGVGRKQADRPQYTLPLAGKTVIEHTLTRLESDFQVAVASVDEPYWPELS 92

QY 69 IAKHPDITTAGGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 128
 DB 93 IAKHPDITTAGGKERADSVLSA-LKALEDIASBNVTVHNAAPCCLTGSNHIHQIDTLK 148

QY 129 NDVGGILASSDHTLKHVDG-TITATIDRKHYVRLTPQWFKYGMALDALQ-TEG 184
 DB 149 TSVGGILAVPADITLKSDRDSRVSITIDRSYVVLATPQWFKYGMALDALQ-TEG 208

QY 185 NPATDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 227
 DB 209 ATITDEASALELGHKPKIVEGRPDNITKTRPBDLALAQ 250

RESULT 7
 US-09-815-242-10273

```

Sequence 10273, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10273
LENGTH: 236
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10273

Query Match          43.4%; Score 520.5; DB 9; Length 236;
Best Local Similarity 49.8%; Pred. No. 2,2e-45;
Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

QY 9 AVPAAGVGRMADPCKOYPLAGKTVIEHTLTRLLSDAPQKVAVAISVEDPYPELS 68
DB 10 AVPAAGVGRMADPCKOYPLAGKTVIEHTLTRLLSDAPQKVAVAISVEDPYPELS 69
QY 69 IAKHPDITAPGKERADSVLSALKLEDIASENDVYLVDARPCITGSDIHLQIDTLK 128
DB 70 LANHPQITVVDGGERADSVLSAGLKA---AGDAQWVLVDARPCILHODDLARLLALS 125
QY 129 NDPVGGILALSHDITLKHVD--GDTITATIDRKVWRALTPQMKYKMLRDLAQ--TEG 184
DB 126 TSRTGGILAAPVDITMKRAEPGKNAIAHTVDKGLMHALTPQFPRELLHDCULRALNG 185
QY 185 NPAYTDEASALELLGHRKIVGSRPPNIKITRPEDIALAQFTM 227
DB 186 -ATTDEASALEVCGFHPQVEGRADNIKVTRPEDIALAEFYL 227

RESULT 8
US-09-792-251-11
Sequence 11, Application US/09792251
Patent No. US20020160364A1
GENERAL INFORMATION:
APPLICANT: Fultz, Christian
APPLICANT: Youngman, Phillip
APPLICANT: Guzman, Luz-Maria
APPLICANT: Guzman, Luz-Maria
TITLE OF INVENTION: USE OF YACM AND YQD1, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES
FILE REFERENCE: 06286-140001
CURRENT APPLICATION NUMBER: US/09/792,251
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 28

```

```

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 236
TYPE: PRT
ORGANISM: Escherichia coli
US-9-792-251-11

Query Match          43.4%; Score 520.5; DB 10; Length 236;
Best Local Similarity 49.8%; Pred. No. 2,2e-45;
Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

QY 9 AVPAAGVGRMADPCKOYPLAGKTVIEHTLTRLLSDAPQKVAVAISVEDPYPELS 68
DB 10 AVPAAGVGRMADPCKOYPLAGKTVIEHTLTRLLSDAPQKVAVAISVEDPYPELS 69
QY 69 IAKHPDITAPGKERADSVLSALKLEDIASENDVYLVDARPCITGSDIHLQIDTLK 128
DB 70 LANHPQITVVDGGERADSVLSAGLKA---AGDAQWVLVDARPCILHODDLARLLALS 125
QY 129 NDPVGGILALSHDITLKHVD--GDTITATIDRKVWRALTPQMKYKMLRDLAQ--TEG 184
DB 126 TSRTGGILAAPVDITMKRAEPGKNAIAHTVDKGLMHALTPQFPRELLHDCULRALNG 185
QY 185 NPAYTDEASALELLGHRKIVGSRPPNIKITRPEDIALAQFTM 227
DB 186 -ATTDEASALEVCGFHPQVEGRADNIKVTRPEDIALAEFYL 227

RESULT 9
US-9-815-242-13867
Sequence 13867, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13867
LENGTH: 236
TYPE: PRT
ORGANISM: Salmonella typhi
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(236)
OTHER INFORMATION: Xaa = Any Amino Acid
US-9-815-242-13867

Query Match          41.7%; Score 499.5; DB 9; Length 236;

```

Best Local Similarity 48.4%; Pred. No. 3.2e-43;
Matches 108; Conservative 32; Mismatches 74; Indels 9; Gaps 4;

QY 9 AVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 68
Db 10 AVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 69
QY 69 IAKRPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 128
Db 70 IAKRPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 125
QY 129 NDVPVGLIASSHDTLKVD--GDTITATIDRKHWRAALTPQWFKYGMELDALQR--TEG 184
Db 126 NSRVGGLIASPVRTMRKGEPRKQNAIAHYERADLMALTPQFPRELLMDCLTRALNIG 185
QY 185 NPATVDEASALELGHKPKIVEGRPNIKITRPELDALAQFYM 227
Db 186 ATITDEASALEYCGFHPALVEGRADNIKVTRPELDALAEFTL 227

RESULT 10

US-09-792-251-8
; Sequence 8, Application US/09792251
; Patent No. US20020160364A1
; GENERAL INFORMATION:
; APPLICANT: Fritz, Christian
; APPLICANT: Youngman, Philip
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: USE OF YACM AND YOBJ, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES
; TITLE OF INVENTION: AND THEIR USE
; FILE REFERENCE: 06286-140001
; CURRENT APPLICATION NUMBER: US/09/792.251
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-792-251-8

Query Match 24.8%; Score 297; DB 10; Length 232;
Best Local Similarity 36.8%; Pred. No. 2.5e-22;
Matches 85; Conservative 40; Mismatches 88; Indels 18; Gaps 9;

QY 10 VVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 68
Db 6 VVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 64
QY 69 IAKRPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 125
Db 65 IAKRPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHLQDITLK 116
QY 126 TL--KNDPVG--GILASSHDTLKVDGDTITATIDRKHWRAALTPQWFKYGMELDALQR 182
Db 117 ELIASEGTGAALAVVVKITKRVQDIQVSEITSESSVAVQTPQAFRLSLMKHAA 176
QY 183 EGNPAV--TDEASALEL--GHKPKIVEGRPNIKITRPELDALAQFYMEQQA 231
Db 177 ERKPGITGDSALVEQMGGSVRYVEGSYTNIKLTPDILTSALAEISES 227

RESULT 11

US-09-935-943-8
; Sequence 8, Application US/09935943
; Patent No. US20020120963A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Megrich, Lynette M.
; APPLICANT: Budziszewski, Gregory J.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-31509A
; CURRENT APPLICATION NUMBER: US/09/935.943

; CURRENT FILING DATE: 2001-08-23
; FOR APPLICATION NUMBER: PCT/EP01/08910
; FILING DATE: 2001-08-01
; FOR APPLICATION NUMBER: US 60/222,779
; FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-935-943-8

QY Match 23.6%; Score 263.5; DB 10; Length 302;
Best Local Similarity 34.2%; Pred. No. 9e-21;
Matches 80; Conservative 46; Mismatches 83; Indels 25; Gaps 10;

QY 10 VVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 63
Db 82 ILLAGGQGRKMKMMPKQYIPLLQPIALVSFFPSRMP-----YKAIWVCCDPFRD 135
QY 64 -WPELSTAKHPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHL 122
Db 136 IFEYVESIDVLLPFLPKERQDSVYSGLQEI--DVASE--LYCINHSARFLVYTEVE- 191
QY 123 QIDTLKNDPVGG--ILASSHDTLKVDGDT--ITATIDRKHWRAALTPQWFKYGMELDA- 178
Db 192 --KVLKQSAVGAALVGPAPAKATKEVNSDSLVTKTLDRKLTWEMQPVYKPELKKGF 249
QY 179 -LQRTGKPNATDEASALELGHKPKIVEGRPNIKITRPELDALAQFYMEQQA 231
Db 250 ELVASEG--LEVTDVDSIVELYKHPVVSQSYTNIKVTPDILTLERILSEDS 302

RESULT 12

US-1-128-713A-6
; Sequence 6, Application US/10128713A
; Patent No. US20030170847A1
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Michael G
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production
; FILE REFERENCE: CL-1788
; CURRENT APPLICATION NUMBER: US/10/128,713A
; CURRENT FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis
US-1-128-713A-6

QY Match 21.3%; Score 255.5; DB 12; Length 232;
Best Local Similarity 35.0%; Pred. No. 4.8e-18;
Matches 78; Conservative 30; Mismatches 96; Indels 19; Gaps 5;

QY 9 AVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 67
Db 5 AVPAAGVGMQADRPQYPLAGKTVIEHTLTLLSDFOKVAVAISVEDPYPELS 58
QY 68 -----SIKRPDITTAAGKERADSVLSALKALEDIASENDWVLYHDAARPCLTGSDIHL 122
Db 59 EGVVADLGRASDVVVGGAERTDSVAGLSA---AGDAPVLYHDAARALPPALILAR 114
QY 123 QIDTLKNDPVGG--ILASSHDTLKVDGDT--ITATIDRKHWRAALTPQWFKYGMELDALQR 181
Db 115 VVDAIRAGSSAVLPLVVTITIKSVDLGAVTGPIPLSRRAVQTPQGFSTVLRSAVD- 173
QY 182 TEGNPAVDEASALELGHKPKIVEGRPNIKITRPELDALAQFYMEQQA 224
Db 174 -AGVNAATDADALVERGVSVQITIPDADALAFKITTPLDVLAR 215

RESULT 13
 US-10-156-761-11503
 Sequence 11503, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OKURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIGAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: KATORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 11503
 LENGTH: 250
 TYPE: PR1
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-11503

Query Match 20.5%; Score 245.5; DB 15; Length 250;
 Best Local Similarity 32.1%; Pred. No. 5,7e-17;
 Matches 75; Conservative 34; Mismatches 96; Indels 29; Gaps 6;

QY 9 AVVPAAGVGMQADRPKQYLPLAGKTVIEHTLRLLESAPQKVAAL-----SV 59
 DB 16 AVVPAAGVGMQADRPKQYLPLAGKTVIEHTLRLLESAPQKVAAL-----SV 75
 QY 60 EDVWPELSIAKHPDITAPGKERADSVLSALKALDIASNDWLVHDAAPCCTGSD 119
 DB 76 LDAN-----ALPERDFFVVGSGRSQSVLGLDAL---PPGIDIVLVHDAAPCCTGSD 128
 QY 120 IHLQIDTLKNDPVGGIILSSHTLKHV-----DGDITATIDKRVHRAALTPQMFXY 172
 DB 129 VDVAIEAVRDGAPVAVPALPLADTVKQVEPAVPGEPVAVPEPARIARAVOTPO---- 184
 QY 173 GMLRDALQRTGNGDA--VTDEASALELIGHKPKIVEGRPNIKITREDEALAQ 224
 DB 185 GFDDITLVAHEITVDVTDASVWEQAGARVVVVPCHBEAFKVTPLDIVLAE 238

RESULT 14
 US-10-289-762-621
 Sequence 621, Application US/10289762
 Publication No. US20040006218A1
 GENERAL INFORMATION:
 APPLICANT: Griffiths, R.
 TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 FILE REFERENCE: 9710-003-999
 CURRENT APPLICATION NUMBER: US/10/289,762
 CURRENT FILING DATE: 2003-03-27
 NUMBER OF SEQ ID NOS: 6843
 SEQ ID NO 621
 LENGTH: 218
 TYPE: PR1
 ORGANISM: Chlamydia pneumoniae
 US-10-289-762-621

Query Match 20.1%; Score 240.5; DB 12; Length 218;
 Best Local Similarity 32.9%; Pred. No. 1.5e-16;
 Matches 74; Conservative 41; Mismatches 91; Indels 19; Gaps 9;

QY 3 PTIQCAVAVPAAGVGMQADRPKQYLPLAGKTVIEHTLRLLESAPQKVAALSVEDP 62

DB 7 PWKSLILLSGGQTRGSKIPKQYLPANGPLVHSL-KILSS--LPQIAEVIWVCDP 63
 QY 63 YNPESIANHPDITAPGKERADSVLSALKALDIASNDWLVHDAAPCCTGSDHL 122
 DB 64 SYOE-TFOEYVPSAIP-GERRQDSVFGLOOV-----SYPAVLIHDAARPPYDETH- 115
 QY 123 QIDTLKNDPVGGIILSS--HDTLGHVGDITATIDKRVHRAALTPQMFXYGMLRALQ 180
 DB 116 --DLDTAKIGATLASPIPTIKQNP---VRLDNDNLAIHTPQCIKTEILREGLA 170
 QY 181 -RTGPAVTDASALELIGHKPKIVEGRPNIKITREDEALAQ 224
 DB 171 LAERQTLTVDDIEAAEIIIGKPSQVFNHPOIKISTYEDLTIQ 215

RESULT 15
 US-9-738-626-6422
 Sequence 6422, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENO, AKIHIRO
 APPLICANT: IKEDA, MASARO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159152
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 OFTWARS: PatentIn ver. 3.0
 SEQ ID NO 6422
 LENGTH: 256
 TYPE: PR1
 ORGANISM: Corynebacterium glutamicum
 US-9-738-626-6422

Query Match 18.6%; Score 223.5; DB 10; Length 256;
 Best Local Similarity 30.7%; Pred. No. 1.1e-14;
 Matches 71; Conservative 43; Mismatches 102; Indels 15; Gaps 7;

QY 9 AVVPAAGVGMQADRPKQYLPLAGKTVIEHTLRLLESAPQKVAALSVEDP 67
 DB 10 ALVLAAGGTLAGPITKAFVTLERTLERSIQAMLTSSVDEIILVSPDEYTRADL 69
 QY 68 ----SIAGHP---ITAPGKERADSVLSALKALDIASNDWLVHDAAPCCTGSD 119
 DB 70 LRKGLNDPVGAVRVLVGGGERADSVMAQLAISDQAPPAIVLHDSARALTFPPGM 129
 QY 120 IHLQIDTLKNDPVGGIILSSHTLKHV-----DGDITATIDKRVHRAALTPQMFXYGLRD 177
 DB 130 IARVVRKVEGATVAVIPVLPVSDIKKVSPPGVGVVVTIPRAEIRAVOTPOGGLSELVA 189
 QY 178 ALQR--TSGNPAV--VTDEASALELIGHKPKIVEGRPNIKITREDEALAQ 224
 DB 190 ANEKFPADPNPGFIPPTDASLMEWYGVADVVCQDPKAFKVTTEIDMMLAQ 240

Search completed: January 29, 2004, 16:21:16
 Job time: 16.7685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:40:33 ; Search time 21.153 Seconds
(without alignment)

1736.455 Million cell updates/sec

Title: US-09-941-947a-10

Perfect score: 1199

Sequence: 1 MNPTICQMAVPAAGVGKRM.....IKITRPEDLAAQFVMEQQA 231

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
11: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1199	100.0	231	23	ABG61584
2	1199	100.0	231	23	AAE22303
3	1199	100.0	231	23	AAU80327
4	555.5	44.2	225	22	AAU36338
5	530.5	44.2	225	22	AAU35483
6	529.5	44.2	254	22	AAU16129
7	524.5	43.7	236	21	AAU47964
8	520.5	43.4	236	21	AAU13772
9	520.5	43.4	236	22	AAU34680

1	520.5	43.4	235	24	AAE31678	Escherichia coli Y
1	499.5	41.7	236	22	AAU38274	Salmonella typhi C
1	411.5	34.3	229	21	AAV74694	Neisseria meningit
1	408.5	34.1	851	22	ABG16344	Novel human diagno
1	397.5	33.2	229	21	AAV74692	Neisseria gonorrhoe
1	397.5	33.2	229	24	ABP79613	N. gonorrhoeae anti
1	297	24.8	232	24	AAE31677	Bacillus subtilis
1	283.5	23.6	227	22	AAE31678	Arabisopsis thalia
1	283.5	23.6	263	21	AAE31678	Arabisopsis thalia
1	283.5	23.6	276	21	AAE31678	Arabisopsis thalia
2	283.5	23.6	300	21	AAE31678	Arabisopsis thalia
2	283.5	23.6	302	21	AAE31678	Arabisopsis thalia
2	283.5	23.6	302	22	AAE31678	Arabisopsis thalia
2	283.5	23.6	302	22	AAE31678	Arabisopsis thalia
2	283.5	23.6	302	23	AAE31678	Arabisopsis thalia
2	283.5	23.6	322	21	AAE31678	Arabisopsis thalia
2	283.5	23.6	324	21	AAE31678	Arabisopsis thalia
2	283.5	23.6	940	22	ABG30158	Novel human diagno
2	270	22.5	232	23	ABG48960	Listeria monocytog
2	258.5	21.6	211	21	AAE31678	Arabisopsis thalia
3	258.5	21.6	211	21	AAE31678	Arabisopsis thalia
3	258.5	21.6	211	21	AAE31678	Arabisopsis thalia
3	254.5	21.2	207	21	AAE31678	Arabisopsis thalia
3	254.5	21.2	209	21	AAE31678	Arabisopsis thalia
3	254.5	21.2	209	21	AAE31678	Arabisopsis thalia
3	249	20.8	152	21	AAV74693	Neisseria meningit
3	246	20.5	236	23	ABG48961	Listeria monocytog
3	243	20.3	298	23	AAU50683	Propionibacterium
3	240.5	20.1	218	20	AAE31678	Chlamydia pneumonia
4	229.5	19.1	239	23	ABE30406	Streptococcus poly
4	229.5	19.1	247	23	ABE27297	Streptococcus poly
4	224.5	18.7	291	23	ABE54652	Bifidobacterium 10
4	223.5	18.6	256	22	AAE32668	C glutamicum prote
4	223.5	18.6	256	22	AAE32668	Corynebacterium gl
4	223.5	18.6	256	22	AAE32668	Corynebacterium gl

ALIGNMENTS

RESULT 1	ABG6134	3661584 standard; Protein: 231 AA.
ID	3661584	
AC	3661584	
XX	7-AUG-2002 (first entry)	
DE	High growth methanotrophic bacterial strain polypeptide #34.	
XX	High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme;	
XX	ethane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;	
KW	/phosphatase dependent phosphofructokinase; nitrogen-containing compound;	
KW	amonia; nitrate; nitrite; nitrogen; oxygen; landfill;	
KW	ethane-containing environment; waste water treatment system; isoprenoid;	
KW	nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.	
OS	Methylobacter 16a.	
XX		
PN	200220728-AA2.	
XX		
PD	7-MAR-2002.	
XX		
PF	7-AUG-2001; 2001MO-US26827.	
XX		
PR	1-SEP-2000; 2000US-229858P.	
XX		
PA	(UPO) DU PONT DE NEMOURS & CO E. I.	
XX		
PI	Offas M, Odum JM, Schenkele A;	
XX		
DR	1-SEP-2002-452200/48.	

DR N-PSDB; ABK83263.
 XX
 XX New high growth methanotrophic bacterial strain, useful for producing
 PT single cell proteins, grows on a C1 carbon substrate, and comprises a
 PT functional gene encoding in Embden-Meyerhof carbon pathway
 XX
 PS Claim 11: Page 143-144; 157pp; English.

XX The invention relates to a high growth methanotrophic bacterial strain,
 CC which grows on a C1 carbon substrate e.g. methane and methanol, and
 CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
 CC gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16s
 CC RNA. The bacterial strain is useful for the production of single cell
 CC protein and for the bioconversion of a nitrogen-containing compound,
 CC e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
 CC production of a feed product comprising a protein, carbohydrates and a
 CC pigment and for reducing oxygen demand, for removing nitrates and
 CC nitrates in methane-containing environments such as landfills, waste
 CC water treatment systems or anywhere that methane, oxygen and nitrates are
 CC present. The bacterial strain of the invention can be used as a
 CC denitrifying agent for the conversion of nitrate or nitrite to nitrous
 CC oxide with methane or methanol as a carbon source. It is also used in the
 CC production of biomass including proteins, carbohydrates and a wide
 CC variety of pigments (particularly for isoprenoid pigments for the
 CC purpose of generating animal feeds), in production of terpeneoid and
 CC carotenoid compounds, useful as pigments and as monomers in polymeric
 CC materials and in production of exopolysaccharides at high levels.
 CC Sequences ABG61551-ABG61590 represent high growth methanotrophic
 CC bacterial strain proteins of the invention.

XX Sequence 231 AA:
 SQ
 Query Match 100.0%; Score 1199; DB 23; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5,7e-118;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITQCAVAVPAAGVGRMQRKQYLPAGKTVIHTLTRLLIESAPFQXVAVLSVE 60
 DB 1 MNPITQCAVAVPAAGVGRMQRKQYLPAGKTVIHTLTRLLIESAPFQXVAVLSVE 60
 QY 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120
 DB 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120
 QY 121 HQIDITLKNDDPVGGILAISSHDTLKHVGDITITATIDRKHWRAALTPOMFYGYMLRDALQ 180
 DB 121 HQIDITLKNDDPVGGILAISSHDTLKHVGDITITATIDRKHWRAALTPOMFYGYMLRDALQ 180
 QY 181 RTEGNPAVTVDEASALELIGHKPKIVGRPDNIXITRPBDLALAQFYVEQQA 231
 DB 181 RTEGNPAVTVDEASALELIGHKPKIVGRPDNIXITRPBDLALAQFYVEQQA 231

RESULT 2
 AAEE22303
 ID AAEE22303 standard; Protein; 231 AA.
 XX
 AC AAEE22303;

XX 25-UTL-2002 (first entry)
 DE Methylomonas 16a sp. 2C-methyl-d-erythritol cytidyltransferase enzyme.
 XX
 KM Carotenoid; isopentenyl pyrophosphate; anthraxanthin; astaxanthin; diet;
 KM anti-oxidant; steroid; flavour; fragrance; electro-optic application;
 KM aquaculture; enzyme; 2C-methyl-d-erythritol cytidyltransferase; ispd.
 XX
 OS Methylomonas 16a sp.
 XX
 PN WO00218617-A2.
 XX
 PD 07-MAR-2002.
 XX

PF 04-SEP-2001; 2001MO-US27420.
 XX
 PR 01-SEP-2000; 2000US-229858P.
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPC) DU POINT DE MEMOIRS & CO E I.
 XX
 PI Brzostowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;
 PI Odum JM, Picotaggo SK, Rouviere PR;
 DR WPI; 2002-351711/38.
 DR N-PSDB; MAD35501.

XX Producing carotenoid compounds e.g. anthraxanthin and astaxanthin, by
 PT using microorganisms having a nucleic acid molecule encoding enzymes in
 PT the carotenoid biosynthetic pathway and which metabolize single carbon
 PT substrates
 XX
 PS Claim 43; Page 156pp; English.

XX The invention relates to a method for producing carotenoid compounds.
 CC The method comprises a transformed metabolising host cell, comprising
 CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule
 CC encoding an enzyme in the carotenoid biosynthetic pathway, under the
 CC control of regulatory sequences, and contacting the host cell with carbon
 CC substrate to produce a carotenoid compound. The method is useful for
 CC producing carotenoid compounds such as anthraxanthin and astaxanthin, by
 CC using microorganisms having a nucleic acid molecule encoding enzymes in
 CC the carotenoid biosynthetic pathway and which metabolize single carbon
 CC substrates. The carotenoids have potent anti-oxidant properties useful in
 CC diet, and aquaculture elements. The carotenoids are also useful as
 CC intermediates in the synthesis of steroids, flavours and fragrances and
 CC compounds for potential electro-optic applications. The present sequence
 CC is Methylomonas 16a sp. 2C-methyl-d-erythritol cytidyltransferase (ispd)
 CC enzyme used in the invention.

XX Sequence 231 AA:
 SQ
 Query Match 100.0%; Score 1199; DB 23; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5,7e-118;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITQCAVAVPAAGVGRMQRKQYLPAGKTVIHTLTRLLIESAPFQXVAVLSVE 60
 DB 1 MNPITQCAVAVPAAGVGRMQRKQYLPAGKTVIHTLTRLLIESAPFQXVAVLSVE 60
 QY 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120
 DB 61 DYPWPELSIAKHPDITITAPGKERADSVLSALKALEDIASENDWLVVDARPCITGSDI 120
 QY 121 HQIDITLKNDDPVGGILAISSHDTLKHVGDITITATIDRKHWRAALTPOMFYGYMLRDALQ 180
 DB 121 HQIDITLKNDDPVGGILAISSHDTLKHVGDITITATIDRKHWRAALTPOMFYGYMLRDALQ 180
 QY 181 RTEGNPAVTVDEASALELIGHKPKIVGRPDNIXITRPBDLALAQFYVEQQA 231
 DB 181 RTEGNPAVTVDEASALELIGHKPKIVGRPDNIXITRPBDLALAQFYVEQQA 231

RESULT 3
 AAUE 327
 ID AAUE327 standard; Protein; 231 AA.
 XX
 AC AAUE327;
 XX
 DE 15-UTL-2002 (first entry)
 DE Methylomonas 16a ORF3 ysbp/ispd protein sequence.
 XX
 KM Isoprenoid biosynthetic enzyme; isoprenoid compound; feed additive;
 KM carotenoid; pigment; flavour; fragrance; open reading frame 3; ORF3;
 KM ysbp; 2C-methyl-D-erythritol cytidyltransferase enzyme; ispd.
 XX

OS Methylomonas sp.
 XX WO200220733-A2.
 XX
 XX 14-MAR-2002.
 XX
 XX 29-AUG-2001; 2001WO-US26852.
 XX
 XX 01-SEP-2000; 2000US-2299072.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 XX Cheng Q, Koffas M, Norton KC, Odom JM, Picataggio SK, Rouviere PE,
 XX Schenzle A, Tomb J,
 XX WPI: 2002-383051/41.
 XX N-PSDB; AAKS0083.
 XX
 XX Novel nucleic acid molecule encoding a isoprenoid biosynthetic enzyme,
 XX isolated from Methylomonas 16a, useful for the production of isoprenoid
 XX compounds -
 XX
 XX Claim 4; Page 70-71; 84pp; English.
 XX
 XX The present invention relates to a new nucleic acid molecule encoding
 XX an isoprenoid biosynthetic enzyme isolated from Methylomonas 16a.
 XX The invention is useful for obtaining a nucleic acid molecule
 XX encoding an isoprenoid compound biosynthetic enzyme, and for the
 XX microbial production of isoprenoid compounds. The molecules of the
 XX invention are also useful for regulating isoprenoid biosynthesis in an
 XX organism and for producing recombinant organisms for producing various
 XX isoprenoid compounds. The nucleic acid is also useful for feed additive,
 XX for the production of kerateneoids and their derivatives, isoprenoid
 XX intermediates, and as pure products useful as pigments, flavours and
 XX fragrances. The present amino acid sequence represents the Methylomonas
 XX 16a open reading frame 3 (ORF3) ygbP/ispD (2C-methyl-D-erythritol
 XX cytidyltransferase enzyme) protein of the invention, as described above.
 XX
 XX Sequence 231 AA;
 XX
 XX Query Match 100.0%; Score 1199; DB 23; Length 231;
 XX Best Local Similarity 100.0%; Pred. No. 5,76-118;
 XX Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 MNPTTQCAVAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSSDAFOKVAVALSV 60
 XX 1 MNPTTQCAVAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSSDAFOKVAVALSV 60
 XX
 XX 61 DPYAPDELSTAKHPDITTAAGGERADSVLSALKLEDIASENDVWLVHDAARPCLTGSDI 120
 XX 61 DPYAPDELSTAKHPDITTAAGGERADSVLSALKLEDIASENDVWLVHDAARPCLTGSDI 120
 XX
 XX 121 HQIDTTLKNDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 180
 XX 121 HQIDTTLKNDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 180
 XX
 XX 181 RTEGNPAVTDASALELGHKPKIVEGRPDNIKITRPEDLALAQFYMEQQA 231
 XX 181 RTEGNPAVTDASALELGHKPKIVEGRPDNIKITRPEDLALAQFYMEQQA 231
 XX
 XX RESULT 4
 XX AAU36338 standard; Protein; 234 AA.
 XX
 XX AAU36338;
 XX
 XX 14-FEB-2002 (first entry)
 XX
 XX Pseudomonas aeruginosa cellular proliferation protein #328.
 XX
 XX Antisense; prokaryotic cellular proliferation protein;
 XX antibiotic; antibacterial; drug design.
 XX

XX
 XX OS Pseudomonas aeruginosa.
 XX
 XX PN 0200170955-A2.
 XX
 XX XX 07-SEP-2001.
 XX
 XX PD 1-MAR-2001; 2001WO-US09180.
 XX
 XX XX 1-MAR-2000; 2000US-191078P.
 XX
 XX XX 3-MAY-2000; 2000US-206848P.
 XX
 XX XX 5-MAY-2000; 2000US-207727P.
 XX
 XX XX 3-OCT-2000; 2000US-242578P.
 XX
 XX XX 7-NOV-2000; 2000US-253625P.
 XX
 XX XX 3-DEC-2000; 2000US-257931P.
 XX
 XX XX 5-FEB-2001; 2001US-269308P.
 XX
 XX XX ELIT-1) ELITRA PHARM INC.
 XX
 XX PI Maselbeck R, Ohlisen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 XX PI Masamoto RT, Xu HH;
 XX
 XX XX P1; 2001-611495/70.
 XX
 XX DR 1-PSDB; AAS54197.
 XX
 XX XX 1-PSDB; AAS54197.
 XX
 XX PS New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 11931; 511pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 XX prokaryotic cellular proliferation, their use in identifying the
 XX genes, their use in the discovery of novel antibiotics, the essential
 XX genes themselves and the encoded proteins. The prokaryotes used are
 XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 XX invention is also useful for the identification of potential new targets
 XX for antibiotic development. The antisense nucleic acids can also be used
 XX to identify proteins used in proliferation, to express these proteins,
 XX and to obtain antibodies capable of binding to the expressed proteins.
 XX The proteins can be used to screen compounds in rational drug discovery
 XX programmes. The antisense nucleic acid sequence is also useful to screen
 XX for homologous nucleic acids which are required for cell proliferation in
 XX a wide variety of organisms. The present sequence represents an
 XX essential prokaryotic cellular proliferation protein.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX http://wipo.int/pub/published_pct_sequences.
 XX
 XX XX Sequence 234 AA;
 XX
 XX Query Match 46.3%; Score 555.5; DB 22; Length 234;
 XX Best Local Similarity 54.1%; Pred. No. 4,98-50;
 XX Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;
 XX
 XX 8 NAAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSSDAFOKVAVALSV 67
 XX 8 NAAVPAAGVGKMQADRPKQYLPLAGKTVIEHTYTRLLSSDAFOKVAVALSV 67
 XX
 XX 10 WTVTPAAGVGSRMRDRPKQYLDLGRVIEHTYTRLLSSDAFOKVAVALSV 69
 XX 10 WTVTPAAGVGSRMRDRPKQYLDLGRVIEHTYTRLLSSDAFOKVAVALSV 69
 XX
 XX 68 STAKHPDITTAAGGERADSVLSALKLEDIASENDVWLVHDAARPCLTGSDI 126
 XX 68 STAKHPDITTAAGGERADSVLSALKLEDIASENDVWLVHDAARPCLTGSDI 126
 XX
 XX DB 70 DCAARHVRQRAAGGAGGSRGVTNLGLRLEGAQADQWLVHDAARNTRGDRLLE 129
 XX 70 DCAARHVRQRAAGGAGGSRGVTNLGLRLEGAQADQWLVHDAARNTRGDRLLE 129
 XX
 XX DB 127 LKNDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 184
 XX 127 LKNDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 184
 XX
 XX DB 130 LAEDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 189
 XX 130 LAEDPVGGIALLSSHDITLKHVGDITITATIDRKHWYRALTPQMFYGMRLDALQ 189
 XX
 XX DB 185 NPAAVTDASALELGHKPKIVEGRPDNIKITRPEDLALAQFYMEQQA 224
 XX 185 NPAAVTDASALELGHKPKIVEGRPDNIKITRPEDLALAQFYMEQQA 224
 XX
 XX DB 190 GVAITDEASAMENAGYAPKIVEGRADNLIKITRPEDLALQ 229
 XX 190 GVAITDEASAMENAGYAPKIVEGRADNLIKITRPEDLALQ 229
 XX

RESULT 5
AAU35483
ID AAU35483 standard; Protein; 225 AA.
XX
XX AAU35483;
AC
XX
XX 14-FEB-2002 (first entry)
DT
XX
XX Haemophilus influenzae cellular proliferation protein #124.
DE
XX
XX Actisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
XX Haemophilus influenzae.
OS
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
DR N-PSDB; AAS53342.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11076; 511pp; English.
PS
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 225 AA;
SQ

Query Match 44.2%; Score 530.5; DB 22; Length 225;
Best Local Similarity 52.5%; Pred. No. 2e-47;
Matches 117; Conservative 28; Mismatches 65; Indels 13; Gaps 6;

QY 9 AVPPAGVGKMQADRPQYVPLAGKTVIEHTLRLSSDAFORVAVAISVEDEPYWELLS 68
DB 7 AVLPAAGVSRMQADKPKQYVTLTGKTLIEHTLQVMSYPAVSKITLAVSDPDYISTLS 66
QY 69 IAKHDIITABGGRKADSVLSALKALDIASNDWVYVHNAARPCLTGSDIHQIDTLK 128

DB 67 L--DEKIQVSGGTTIRASVANGINA---LAKNAWVYVHNAARPCLO---HADIDKL 117
QY 129 --NDEVGIIALSSHTLKHVDG--DTYATIDRKHWRALTPQFKKGMALDALQR--TEG 184
DB 118 AIEDKQGAIIAIPVDTIKRADNOQCIKVEDRSQIWMQAMTPQFFPDIIIDALSTGICQ 177
QY 185 NPAVTDEASALELIGHXKPKIVEGRPDNKKIRPDLLAOPYM 227
DB 178 GANITDEASALFELAGFRPHLVGRSDNKLKVRPDLALAEVYL 220

RESULT 6
AAU3129
ID AAU3129 standard; Protein; 254 AA.
XX
XX AAU3129;
AC
XX
XX 14-FEB-2002 (first entry)
DT
XX
XX *Klebsiella pneumoniae* cellular proliferation protein #117.
DE
XX
XX Actisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
XX *Klebsiella pneumoniae*.
OS
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
DR N-PSDB; AAS53388.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11722; 511pp; English.
PS
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 254 AA;
 SQ Query Match 44.2%; Score 529.5; DB 22; Length 254;
 Best Local Similarity 51.1%; Pred. No. 3e-47;
 Matches 114; Conservative 30; Mismatches 70; Indels 9; Gaps 5;

QY 9 AVVPAAGVGMQMDRPRQYPLAGKTYIEHTLRLLESDFQKVAVAISVEDEYWEELS 68
 DB 33 AVVPAAGGRMRQTECPKQYLSIGNKTLLEHVAVALADANVQVVAIVBGRFRFQQLP 92
 QY 69 IAKHPDITAPGKGRADSVLSALKALEDIASNDWYVHDARPCTGSDIHLQIDTLK 128
 DB 93 LAQHPQITVVDGGRADSVLSAGLQAL----PEAQWYVHDARPCLHODILSLSLCS 148
 QY 129 NDPVGGILALSHDLTKHVD--GDTITATIDRKVWRALTPQMKYKGLADALQR--TEG 184
 DB 149 TSVRGGILAAPVDTKGABFGKTAIAHTVDRNDIMHAIPTQLPPELHVDCLTRALNEG 208
 QY 185 NPAYTDEASALELLGKPKYIEGRPDNIXITRPEDLALAOFYM 227
 DB 209 -AATIDEASALEYCGFHPQLVAGRADNIXITRPEDLALAEFYL 250

RESULT 7

AA47964
 ID AA47964 standard; Protein; 236 AA.

XX AC AA47964;

DT 05-MAR-2002 (first entry)

DE Escherichia coli CDP-ME synthase.

XX Escherichia coli; YGDP; CDP-ME synthase; protein coordinate data;

KM 4-diphosphocytidylyl-3-C-methylerythritol synthase; terpenoid; infection;

KM non-mevalonate isoprenoid; biosynthesis pathway; antibacterial; tetanus;

KM anti-infective; anti-inflammatory; tuberculosic; Streptococcus; anthrax;

XX toxic shock syndrome; meningitis; gonorrhea; gastroenteritis.

XX Escherichia coli.

XX MO260183769-A2.

XX 08-NOV-2001.

XX 03-MAY-2001; 2001WO-US14371.

XX 03-MAY-2000; 2000US-201589P.

XX 12-DEC-2000; 2000US-255086P.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Noe1 JP, Bowman ME, Richard S;

XX WPI; 2002-089742/12.

XX N-PSDB; ABA05928.

XX Composition, useful for treating bacterial infections and for

XX identifying modulator compounds, comprise crystalline

XX 4-diphosphocytidylyl-2-C-methylerythritol synthase

XX Example 1; Page -; 176pp; English.

XX The invention relates to a composition (I) comprising CDP-ME,

XX 4-diphosphocytidylyl-2-C-methylerythritol synthase in crystalline form. The

XX invention also discloses screening for compounds (II) that inhibit the

XX non-mevalonate isoprenoid biosynthesis pathway. (II) has antibacterial,

XX anti-infective, anti-inflammatory and tuberculosic activity. (II) is

XX useful for inhibiting in vitro or in vivo, the activity of one or more

XX enzymes in the non-mevalonate isoprenoid biosynthesis pathway, in a cell

XX or cell-free environment, and thus modulating the growth of a cell e.g.

XX bacterial cell. (II) is also useful for inhibiting bacterial terpenoid

CC synthesis and treating a subject suffering from a bacterial infection

CC e.g. infection by Streptococcus or Escherichia coli. (II) is also useful

CC for treating disorders caused by bacterial infections, including

CC diphtheria, pneumonia, dysentery, anthrax, rheumatic fever, toxic shock

CC syndrome, mastitis, meningitis, gonorrhea, typhoid fever,

CC asteroenteritis, brucellosis, cholera, bubonic plague, tetanus,

CC tuberculosis and Lyme disease. The present sequence is that of the E.

CC coli CDP-ME synthase.

CC Note: The present sequence is not given in the specification but was

CC generated from the encoding gene sequence at Genbank Accession Number

CC U11362858.

XX SQ Sequence 236 AA;

XX Query Match 43.7%; Score 524.5; DB 23; Length 236;

XX Best Local Similarity 50.2%; Pred. No. 9.2e-47;

XX Matches 112; Conservative 31; Mismatches 71; Indels 9; Gaps 4;

QY 9 AVVPAAGVGMQMDRPRQYPLAGKTYIEHTLRLLESDFQKVAVAISVEDEYWEELS 68

DB 10 AVVPAAGGRMRQTECPKQYLSIGNKTLLEHVAVALADANVQVVAIVBGRFRFQQLP 69

QY 69 IAKHPDITAPGKGRADSVLSALKALEDIASNDWYVHDARPCTGSDIHLQIDTLK 128

DB 70 IAKHPDITVVDGGRADSVLSAGLQAL----AGQWYVHDARPCLHODILSLSLCS 125

QY 129 NDPVGGILALSHDLTKHVD--GDTITATIDRKVWRALTPQMKYKGLADALQR--TEG 184

DB 126 TSVRGGILAAPVDTKGABFGKTAIAHTVDRNDIMHAIPTQLPPELHVDCLTRALNEG 185

QY 185 NPAYTDEASALELLGKPKYIEGRPDNIXITRPEDLALAOFYM 227

DB 186 -AATIDEASALEYCGFHPQLVAGRADNIXITRPEDLALAEFYL 227

RESULT 8

AA41172
 ID AA41172 standard; protein; 236 AA.

XX AC AA41172;

DT 02-FEB-2001 (first entry)

DE Escherichia coli YGDP protein.

XX YGDP; KDTB; YGCF, YGCF, YHBC; YGAP; YGAB; YCHB; antibacterial;

XX treatment; infection.

XX Escherichia coli.

XX U19916176-A1.

XX 02-OCT-2000.

XX 02-APR-1999; 99DE-1016176.

XX 02-APR-1999; 99DB-1016176.

XX (BAYER) BAYER AG.

XX Huetz H, Ehlert K, Freiberg C, Spaltmann F, Wieland B;

XX Labischinski H;

XX WPI; 2000-639611/62.

XX N-PSDB; AAC66042.

XX Essential genes from bacteria, useful in screening for antimicrobial

XX agents, and related proteins, transformants and antisense sequences

XX Enclosure; Page 21-22; 28pp; German.

XX The invention describes novel Escherichia coli genes (I) encoding

XX proteins (II) designated YGCF, YHBC, YGCF, YGAB, YCHB, YGAB, YGAB and

CC KDTs, and genes (Ia) that encode orthologous gene products (Iia) in
 CC other microorganisms and which have antibacterial activity. Recombinant
 CC microorganisms in which expression of (I) or (Ia) can be regulated are
 CC used to identify compounds that bind to the gene products, particularly
 CC in affinity selection assays. (II) and (Iia) are used to identify, or
 CC prepare, antibodies and other proteins that bind to the gene products.
 CC Substances that bind to (II) or (Iia) are potentially useful as
 CC antibacterials for treating a wide range of infections in humans and
 CC animals. Sequences antisense to (I) and (Ia) can also be used as
 CC antibacterials. The specified genes are widely distributed in bacteria
 CC but have no close homologs in eukaryotic cells.

SQ Sequence 236 AA;

Query Match 43.4%; Score 520.5; DB 21; Length 236;

Best Local Similarity 49.8%; Pred. No. 2,4e-46; Mismatches 72; Indels 9; Gaps 4;

Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

DB 9 AVPAAGVGMKQMDRPPQYVPLAGKVIHTLRLIESDAFQKVAIVSEDPYMEIS 68

DB 10 AVPAAGVGMKQMDRPPQYVPLAGKVIHTLRLIESDAFQKVAIVSEDPYMEIS 69

QY 69 IAKHPDITAPGKERADSVLSAKALEDIASENDVYVADARPCITGSDIHLQIDTLK 128

DB 70 LANHPQITVVDGGERADSVLSAKALEDIASENDVYVADARPCITGSDIHLQIDTLK 125

QY 129 NDPVGGITLSSHDLEKVD--GDTTATIDRKHVWALTPQMKYGMALDQ--TEG 184

DB 126 TSTRGTGILAPVRDTPMRAEGKNAIAHTVDRNGLWALTPQFPRELLDCLTRALNS 185

QY 185 NPATVDEASALELGHKPKIVEGRPDNIXITRBDLALAQFM 227

DB 186 -ATTDEASALELGHKPKIVEGRPDNIXITRBDLALAEYL 227

RESULT 9
 AAU34680 standard; Protein; 236 AA.

AC AAU34680;

DT 14-FEB-2002 (first entry)

DE E. coli cellular proliferation protein #261.

KM Antisense, prokaryotic cellular proliferation protein;

OS antibiotic; antibacterial; drug design.

OS Escherichia coli.

PN WO200170955-A2.

PD 27-SEP-2001.

PP 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207227P.

PR 23-OCT-2000; 2000US-243578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX DR N-PSDB; AAS52539.

XX PT New polynucleotides for the identification and development of

PR antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 10273; 51bp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pcr_sequences.

SQ Sequence 236 AA;

QY Match 43.4%; Score 520.5; DB 22; Length 236;

Best Local Similarity 49.8%; Pred. No. 2,4e-46; Mismatches 72; Indels 9; Gaps 4;

Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;

DB 9 AVPAAGVGMKQMDRPPQYVPLAGKVIHTLRLIESDAFQKVAIVSEDPYMEIS 68

DB 10 AVPAAGVGMKQMDRPPQYVPLAGKVIHTLRLIESDAFQKVAIVSEDPYMEIS 69

QY 69 IAKHPDITAPGKERADSVLSAKALEDIASENDVYVADARPCITGSDIHLQIDTLK 128

DB 70 LANHPQITVVDGGERADSVLSAKALEDIASENDVYVADARPCITGSDIHLQIDTLK 125

QY 129 NDPVGGITLSSHDLEKVD--GDTTATIDRKHVWALTPQMKYGMALDQ--TEG 184

DB 126 TSTRGTGILAPVRDTPMRAEGKNAIAHTVDRNGLWALTPQFPRELLDCLTRALNS 185

QY 185 NPATVDEASALELGHKPKIVEGRPDNIXITRBDLALAQFM 227

DB 186 -ATTDEASALELGHKPKIVEGRPDNIXITRBDLALAEYL 227

RESULT 10
 AAB3 578 standard; Protein; 236 AA.

AC AAB31678;

DT 07-MAR-2003 (first entry)

DE Escherichia coli ybp protein.

KM 3-year protein; 5-year protein; pharmaceutical formulation;

OS bacterial infection; antibacterial.

OS Escherichia coli.

PN WO200281652-A2.

PD 17-OCT-2002.

PP 21-FEB-2002; 2002WO-US05086.

PR 23-FEB-2001; 2001US-0792251.

PR (MILL-) MILLENNIUM PHARM INC.

PA Fritz C, Youngman P, Guzman L;

XX MP1; 2003-058529/05.
 DR N-PSDB; AAD48778.
 XX Method for determining whether a test compound is a candidate
 PT antibacterial compound by its effect on the polypeptides encoded by the
 PT genes yacM and S-ygeU
 XX
 PS Disclosure: Fig 4; 49pp; English.
 CC The invention relates to a method for determining whether a test compound
 CC is a candidate antibacterial compound. The method comprising: contacting
 CC an S-yacM or an S-ygeU polypeptide with the test compound; and detecting
 CC interaction of the test compound with the S-yacM or S-ygeU polypeptide,
 CC where an interaction indicates that the test compound is a candidate
 CC antibacterial compound. A method is claimed for treating a bacterial
 CC infection in an organism by administering a therapeutically effective
 CC amount of a pharmaceutical formulation and where the bacterial infection
 CC is a Streptococcus infection. An antibacterial agent can also be
 CC administered to treat a bacterial infection in an organism. The present
 CC sequence is Escherichia coli ygeU protein.
 XX
 SQ Sequence 236 AA;
 Query Match 43.4%; Score 520.5; DB 24; Length 236;
 Best Local Similarity 49.8%; Pred. No. 2,4e-46;
 Matches 111; Conservative 31; Mismatches 72; Indels 9; Gaps 4;
 QY 9 AVPAAGVGKGMQDRKQYPLAGKTVIEHTLRLLESDFGKAVAVISVEDPYPELS 68
 DB 10 AVPAAGGRRMGTCECKOYLSIGNCTILSHVALLAHPRVRVVAISPGDSRPAQLP 69
 QY 69 IAKHPDITAPGKERADSVLSAKALEDIASENDWLVHDAARPCLTGSDIHLQIDTLK 128
 DB 70 LANHPQITVVDGSDERADSVLAGLKA---AGDQWLVHDAARPCLTGSDIHLQIDTLK 125
 QY 129 NDPVGGIILASSHDTLKHVD--GDTITATIDRKHWRAALTPQWFKYGMALDIALOR--TEG 184
 DB 126 TSTRGGIILAPVRDTMRGEPGKNAIAHTVDRNGLMWALTPOFPFRLHLDCTLRALNEG 185
 QY 185 NPAVTDEASALELGHKPKIVEGRPDNIKITRPEDLALAQPYM 227
 DB 186 -ATTIDBASALEYCGFHPALVSGRADNIKVTRPEDLALAEFYL 227
 RESULT 11
 AAU38274
 ID AAU38274 standard; Protein: 236 AA.
 AC AAU38274;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Salmomella typhi cellular proliferation protein #165.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Salmomella typhi.
 XX
 PN MO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.

XX
 PA ELIT-1) ELITRA PHARM INC.
 XX
 PI Iselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI mameto RT, Xu HH;
 XX
 DR 121, 2001-611495/70.
 DR 1-PSDB; AAS56133.
 XX
 PT 2 polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13867; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmomella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC http://wipo.int/pub/published_poc_sequences.
 XX
 SQ Sequence 236 AA;
 Query Match 41.7%; Score 499.5; DB 22; Length 236;
 Best Local Similarity 48.4%; Pred. No. 4e-44;
 Matches 108; Conservative 32; Mismatches 74; Indels 9; Gaps 4;
 QY 9 AVPAAGVGKGMQDRKQYPLAGKTVIEHTLRLLESDFGKAVAVISVEDPYPELS 68
 DB 10 AVPAAGGRRMGTCECKOYLSIGNCTILSHVALLAHPRVRVVAISPGDSRPAQLP 69
 QY 69 IAKHPDITAPGKERADSVLSAKALEDIASENDWLVHDAARPCLTGSDIHLQIDTLK 128
 DB 70 LANHPQITVVDGSDERADSVLAGLQAV---AKQWLVHDAARPCLTGSDIHLQIDTLK 125
 QY 129 NDPVGGIILASSHDTLKHVD--GDTITATIDRKHWRAALTPQWFKYGMALDIALOR--TEG 184
 DB 126 NSRVGGIILASPRDTMRGEPGKNAIAHTVDRNGLMWALTPOFPFRLHLDCTLRALNEG 185
 QY 185 NPAVTDEASALELGHKPKIVEGRPDNIKITRPEDLALAQPYM 227
 DB 186 -ATTIDBASALEYCGFHPALVSGRADNIKVTRPEDLALAEFYL 227
 RESULT 12
 AAV7434
 ID AAV7434 standard; Protein: 229 AA.
 AC AAV7434;
 XX
 DT 11-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 233 protein sequence SEQ ID NO:862.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX

PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 XX 30-APR-1999; 59WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094865.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Mastignani V, Mora M;
 PI Petersen J, Pizzi M, Rapuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR MPI; 2000-062150/05.
 DR N-PSDB; AAZ53456.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 2; Page 534; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ54253 to AAZ55941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunologic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 SQ Sequence 229 AA;
 Query Match 34.3%; Score 411.5; DB 21; Length 229;
 Best Local Similarity 39.7%; Pred. No. 7.4e-35;
 Matches 91; Conservative 48; Mismatches 73; Indels 17; Gaps 5;
 QY 9 AVVPAAGVGRKQADRPQVYLPAGKTYIEHTLRLLESDAFQKVAVAISVEDPY----- 63
 Db 7 ALIPAGIGARFGADKPKQVYVIGSKTYLHTHTGIFERHRAVDLTVVVSPEDFADKVV 66
 QY 64 --WPEHSLAKHPDITTAPEGKRAADSVLSAL-KALE-DIASENDWVYVHDAARCLTGS 119
 Db 67 TAPPOVRYWKN-----GGQTRASTVTRNGVAKLETGTLAETDITLVHDAARCLPSEA 119
 QY 120 IHLQIDTLNDPVGGLILASSHDTLKHVDGDTITATIDRKHWRALTPQMEKYGMRLDAL 179
 Db 120 LRLLEFGAQNAAEGGILAI PVADTLKCADGGMISATVETSTLMQOTQLFRAGILRLAL 179
 QY 180 QRTGKNPAVTDEASALELGHKPKYVEGRPDNFKITRPEDLALAFVME 228
 Db 180 -AAENLDGITDEASAEKGIKIRPLVQGDARNEKLTQDDAVYIVLLD 227

DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #16335.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX Food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dymnac RT, Liu C, Tang YT;
 XX
 DR MPI; 2001-639362/73.
 DR N-PSDB; AAS80531.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 46703; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 851 AA;
 Query Match 34.1%; Score 408.5; DB 22; Length 851;
 Best Local Similarity 46.7%; Pred. No. 1.1e-33;
 Matches 91; Conservative 26; Mismatches 69; Indels 9; Gaps 4;
 QY 9 AVVPAAGVGRKQADRPQVYLPAGKTYIEHTLRLLESDAFQKVAVAISVEDPYWPELS 68
 Db 383 AVVPAAGFRKQVTSRCPQVYLSIGNQITLERSVHALHPVKKVVAISGDSERFALP 442
 QY 69 TAKHPDITTAPEGKRAADSVLSALKALDIASENDWVYVHDAARCLTGSNHLQIDTLK 128
 Db 443 LANHPQITVVDGSDERADSVSLAGLA---AGDAQWVYVHDAARCLTGSNHLQIDTLK 498
 QY 129 NDPVGGIILASSHDTLKHVD--GDTITATIDRKHWRALTPQMEKYGMRLDALQ--TEG 184
 Db 499 TGRGTGILAAPVROTKKAERPKNAIAHTVDRNGMLWALTQGFPRRLHDCITRLALNEG 558
 QY 185 NPATVDASALSLIG 199

RESULT 13
 ABG16344
 ID ABG16344 standard; Protein; 851 AA.
 XX
 AC ABG16344;
 XX

Thu Jan 29 17:41:17 2004

us-09-941-947a-10, tag

Page 10

Db 120 LAALIEBAGAAAGCGILANVPADTLKRAESGGSITATVRSGSGMAQAPQLFQAGLLRL 179

Qy 180 QRTGEGNPAVTFDEASALELLGHKPKIVEGRPNIKITRPEDLALAGFYW 227

Db 180 -AAENLGGITDEASAVEKIGVPRPLLIGQDARNLKLTQPDADIVATLL 226

Search completed: January 29, 2004, 15:49:40
Job time : 22.1153 secs

Thu Jan 29 17:41:17 2004

us-09-941-947a-10.txt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:45:34 : Search time 7.2499 Seconds
(without alignments) 1348.130 Million cell updates/sec

Title: US-09-941-947a-10

Perfect score: 1199

Sequence: 1 MNPTICQMAVPAAGVAGKEM.....IKITREPDIALAQFYEEQQA 231

Scoring table:

ELCOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiletest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	555.5	46.3	270	4	US-09-252-991A-21226 Sequence 21226, A
2	414.5	34.6	242	4	US-09-328-352-7400 Sequence 7400, App
3	240.5	20.1	218	4	US-09-198-452A-621 Sequence 621, App
4	237.5	19.8	238	4	US-09-107-532A-5165 Sequence 5165, App
5	218.5	18.2	271	4	US-09-134-001C-5299 Sequence 2, App1
6	208	17.3	235	3	US-09-285-576-2 Sequence 2, App1
7	104	8.7	461	4	US-09-328-352-5107 Sequence 5107, App
8	93	7.8	357	4	US-09-198-452A-487 Sequence 487, App
9	92	7.5	393	4	US-09-107-532A-6010 Sequence 6010, App
10	89.5	7.5	285	4	US-08-311-731A-34 Sequence 34, App1
11	89.5	7.5	285	4	US-08-311-731A-36 Sequence 36, App1
12	89	7.4	154	4	US-08-936-165A-528 Sequence 528, App
13	89	7.4	531	4	US-09-134-001C-5574 Sequence 3574, App
14	88.5	7.4	459	3	US-08-971-782-2 Sequence 2, App1
15	88.5	7.4	459	3	US-09-109-026-2 Sequence 2, App1
16	87.5	7.3	989	4	US-09-252-991A-17435 Sequence 17435, A
17	86.5	7.2	1788	2	US-08-962-284-2 Sequence 2, App1
18	83.5	7.0	398	3	US-09-303-064-54 Sequence 54, App1
19	83.5	7.0	398	4	US-09-086-503-54 Sequence 54, App1
20	83.5	7.0	405	4	US-09-252-991A-24036 Sequence 24036, A
21	83	6.9	342	3	US-08-867-611-37 Sequence 37, App1
22	83	6.9	344	3	US-08-867-611-38 Sequence 38, App1
23	83	6.9	375	3	US-08-867-611-50 Sequence 50, App1
24	83	6.9	376	1	US-08-500-222-6 Sequence 6, App1
25	83	6.9	376	1	US-08-500-125-6 Sequence 6, App1
26	83	6.9	376	2	US-07-779-704B-6 Sequence 6, App1
27	83	6.9	387	3	US-08-967-611-12 Sequence 12, App1

2	83	6.9	387	5	PCT-US92-06965A-17	Sequence 17, App1
2	83	6.9	393	3	US-08-867-611-14	Sequence 14, App1
3	83	6.9	393	5	PCT-US92-06965A-19	Sequence 19, App1
3	83	6.9	396	3	US-08-867-611-2	Sequence 2, App1
3	83	6.9	396	5	PCT-US92-06965A-7	Sequence 7, App1
3	83	6.9	417	3	US-08-867-611-20	Sequence 20, App1
3	83	6.9	417	5	PCT-US92-06965A-25	Sequence 25, App1
3	83	6.9	425	5	US-08-867-611-22	Sequence 22, App1
3	83	6.9	425	5	PCT-US92-06965A-27	Sequence 27, App1
3	83	6.9	467	3	US-08-867-611-24	Sequence 24, App1
3	83	6.9	467	5	PCT-US92-06965A-29	Sequence 29, App1
3	83	6.9	474	3	US-08-867-611-26	Sequence 26, App1
4	83	6.9	474	5	PCT-US92-06965A-31	Sequence 31, App1
4	83	6.9	491	2	US-08-912-129A-56	Sequence 56, App1
4	83	6.9	496	3	US-08-867-611-10	Sequence 10, App1
4	83	6.9	496	5	PCT-US92-06965A-15	Sequence 15, App1
4	83	6.9	498	1	US-08-500-222-2	Sequence 2, App1
4	83	6.9	498	1	US-08-500-125-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-09-252-991A-21226
Sequence 21226, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUSINOSA FOR DIAGNOSTICS AND THERAPEUTICS
PICTURE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252, 991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074, 788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094, 190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21226
LENGTH: 270
T SE: PRT
I SE: PRT
US-09-252-991A-21226
US-09-252-991A-21226
Pseudomonas aeruginosa

Query Match 46.3%; Score 555.5; DB 4; Length 270;
Best Local Similarity 54.1%; Pred. No. 2.5e-57;
Matches 119; Conservative 26; Mismatches 72; Indels 3; Gaps 3;

QY 8 MAVVPAAGVGRKQMDRPPROYLPKAGTVIEHTTLLIESDAFOKAVAVASVDPYMPDL 67
DB 46 MVTVPAGVGRKQMDRPPROYLPKAGTVIEHTTLLIESDAFOKAVAVASVDPYMPDL 105
QY 68 STAKPDITAPGKERADSVLSA-LKALIEDIASENDVAVDAARPTLTSDDHLDQIT 126
DB 106 DCAASHVGRAGAGRAAGVSVLNGHLLRLEFGAQDDVAVLVDARPTLTGGDDRLIES 165
QY 127 LKNDPVGGLIASSDITLKHVGD-TITATIDRKHVAVLTPQWQKGMALD-LQRTGS 184
DB 166 LKNDPVGGLIASSDITLKHVGD-TITATIDRKHVAVLTPQWQKGMALD-LQRTGS 225
QY 185 NPATVDEASALBLGKRPVIEGRPDNLIKITREPDIALAQ 224
DB 226 GVALTDEASAMEMAGVAPLVYGRADNLIKITREPDIALAQ 265
RESULT 2
US-09-328-352-7400
Sequence 7400, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-039A
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 7400
 ; LENGTH: 242
 ; TYPE: PRN
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-7400

Query Match 34.6%; Score 414.5; DB 4; Length 242;
 Best Local Similarity 43.1%; Pred. No. 1e-40;
 Matches 97; Conservative 34; Mismatches 91; Indels 3; Gaps 3;

QY 8 WAVPAAGVGRMADRPKQYPLAGTVIEHTLRLLESDFQKVAVALSVDPYWEPL 67
 DB 18 WAVPAAGSGSRFSEKTELKQYQDAVLEHVRLSOLPLGTYV-LALGKODTPASTL 76
 QY 68 SIAGHPDITITAPGKERADSVLSALKALBIDASNDVWLVHADAPCLTSGDITLQIDTL 127
 DB 77 SFQDHRKAFHFCNGGVERHVSILNVLNLSQIADSDVWLVHADAPCVTEGCLNTLVKNA 136
 QY 128 KNDPVGGILASSHDTLKHVGD-TITATIDRKAVRALTPQEKYGMALQRT-EGN 185
 DB 137 IETVQSALIAIPVDITLKVNOEQOIKTVSRBLMQAPQPOLAKTILKKAETALKRN 196
 QY 186 PAVTDEASALELGHKPKIVEGRPNIKITSPEDIALAQFMEQO 230
 DB 197 LTIIDFASALSIGHSVQVWGRDNIKITPPDLELAKRLIQSQ 241

RESULT 3

US-09-198-452A-621
 ; Sequence 621, Application US/09198452A
 ; Patent No. 6559294

; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6843
 ; SEQ ID NO 621
 ; LENGTH: 218
 ; TYPE: PRN
 ; ORGANISM: Chlamydia pneumoniae
 US-09-198-452A-621

Query Match 20.1%; Score 240.5; DB 4; Length 218;
 Best Local Similarity 32.9%; Pred. No. 3.6e-20;
 Matches 74; Conservative 41; Mismatches 91; Indels 19; Gaps 9;

QY 3 FTICQWAVPAGVGRMADRPKQYPLAGTVIEHTLRLLESDFQKVAVALSVDP 62
 DB 7 FMKKSILILSGGQCTRFSGSKIPKQYPLNGTPIVLSL-KILSS--LPQIAEVIYVCO 63
 QY 63 YWPELSIAKHDPDITITAPGKERADSVLSALKALBIDASNDVWLVHADAPCLTSGDITL 122
 DB 64 SYQE-TFOEYVPSFAIP-GERRODSVFSGLOV-----SYPMVVIHGGAPFIYDEIH- 115
 QY 123 QIDTFLKNDPVGGILASSHDT-LKHVGDITITATIDRKAVRALTPQEKYGMALQ 180
 DB 116 --DLSTAEKIGATLAAPIPIYTIKQNP---VATIDRDNALIHTTQCKITELHAGLA 170
 QY 181 -RTGEPNAVTTDEASALELGHKPKIVEGRPNIKITSPEDIALAQ 224
 DB 171 LAKEKQTLTVDDIEAAEITIGKRSQIVFNKHPQIKISYPEDITLQ 215

RESULT 4

US-0-107-532A-5165
 ; Sequence 5165, Application US/09107532A

; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660

; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Dereke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5165:
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 218 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium

; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (B) LOCATION 1...238

; SEQUENCE DESCRIPTION: SEQ ID NO: 5165:
 US-0-107-532A-5165

QY Match 19.8%; Score 237.5; DB 4; Length 238;
 Best Local Similarity 28.8%; Pred. No. 9.3e-20;
 Matches 68; Conservative 52; Mismatches 95; Indels 21; Gaps 9;

QY 9 AVPAAGVGRMADRPKQYPLAGTVIEHTLRLLESDFQKVAVALSV--EDPYW-- 64
 DB 9 AIIAGVGKMGODIPQFTTIGKPIIIYTTLSFQHHQVDRIILVCKSGWEETWMA 68
 QY 65 -PELSIAKHDPDITITAPGKERADSVLSALKALBIDASNDVWLVHADAPC--LTSSDI 120
 DB 69 IEFKHSIVKRVIT--GSSKQSHINNGVQFLKQYSDDDTITIIHDIQRLVDELVSIV 126
 QY 121 HLOIDTLKNDPVGGILASSHDT-LKHVGDITITATIDRKAVRALTPQEKYGMALQ 178
 DB 127 IVKCEFGN---GVTSIPVNRQFIKTE-ETTEQYVDENTLIRVSTPOAYGYGKLRRA 181
 QY 179 LQR-TEGNAVTVDA--SALELGHKPKIVEGRPNIKITSPEDIALAQFYEQO 230
 DB 182 YDRAVKEDIGMTSSVYNTWVVDLGTLLYPAAGSDKNIKLTTTDBLELFAIYKMK 237

RESULT 5

US-0-134-001C-5299

```

Sequence 5299, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5299
LENGTH: 271
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5299

```

Query Match	18.2%;	Score 218.5;	DB 4;	Length 271;
Best Local Similarity	29.6%;	Pred. No. 2e-17;		
Matches 72; Conservative	43;	Mismatches 83;	Indels 45;	Gaps 11;

```

QY      8 WAWPAPAGVQXEM-QADRQKQYLPAAGTVEIHTLRLBESDAPQKAAVIAVSEDEYME 66
      36 YALILAGGIGSMWGNPLPKQFLSLQXGPIIHHYKFLMWKXDELIAT-----PQ 88
Db      67 LSTAKHPDII-----TAGGKRADSVLSALKALFD--IASENDWVLVEDANP 113
      89 KWINYMELDINNYQDLDDKKIKYIQQGDNRHNSINNIJESIEHKKUDEDIIVTHDANVP 148
QY      114 CLTGSDI-----HQQIDPLKQ--DPVGGIILATSSHOTLKHVQSDTATATIDKGVW 162
Db      149 FLTNRIIRENWEYASQYGAVDIVNAAVTT-----ISSNDA-QTISGIP-----RSEWY 197
QY      163 RALTPOMKYKGMRLDA-LQRTGNPAV--TDEASALELIGHKPKIVBGRPNIKITPEDL 220
Db      198 QGQTPQTFPKIKELKQSYLSLTQSQKEILTDACKLIVELGKPVKLVKGLFNIKITPYDL 257
QY      221 ALA 223
Db      258 KYA 260

```

RESULT 6
US-09-285-576-2

```

Sequence 2, Application US/09285576
Patent No. 6268177
GENERAL INFORMATION:
APPLICANT: Zalecain, Magdalena
APPLICANT: Burnham, Martin K. R.
APPLICANT: Brown, James R.
APPLICANT: Ingreham, Karen I.
APPLICANT: Chalker, Alison F.
APPLICANT: Holmes, David J.
APPLICANT: Warren, Richard L.
APPLICANT: Mathie, Thomas B.
TITLE OF INVENTION: ncp
FILE REFERENCE: GM10167
CURRENT APPLICATION NUMBER: US/09/285,576
CURRENT FILING DATE: 1999-04-02
EARLIER APPLICATION NUMBER: 60/101,356
EARLIER FILING DATE: 1999-09-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: fastsq for Windows Version 3.0.
SEQ ID NO 2
LENGTH: 235
TYPE: PR1
ORGANISM: Streptococcus pneumoniae
US-09-285-576-2

```

Query Match	17.3%; Score 208; DB 3; Length 235;
-------------	-------------------------------------

Bes Local Similarity 27.6%; Pred. No. 2.8e-16;
 Mismatches 67; Conservative 48; Mismatches 96; Indels 32; Gaps 10;

```

QY      8  WAWVPAAGYKQMO--ADBPXYLFLAGKTYIEHTYLRLESDDAFQKAAAI----- 57
Dh      3  YAGIAGGCTGTMGISNLPKQFELGDRPLIHTIEKFVLEPSIEIKIVGVHGMDWSHA 62

QY      58  SVEDPYWPELSIAKHPIITAPGKEKRAVSLSALKALEDI--ASENDWLVHDAARPC 115
Dh      63  DEVDKILP---LYKERIIIT--KCGADNNTSIKKIIEALVAYRELTPEDIVVTHDSVRPI 118

QY      116  TQSDIHLQDITKAN--DPVGIILASHDLT--KXNDGTTATIDRGVWRAALTPOWEK-- 171
Dh      119  TLRMTQDNIQLAKNNDADVTV--VEADVTTVESTNQFTIDIPNRHLLVQGGTPOTFOR 176

QY      172  -----YGMRLDALQRTGBVNPVDEASALELLGHKPELVGEGEDNKKITRPBDLALAFY 226
Dh      177  DEFDLXGSLSD-----BEKEILTDACKIPIVIGKGDVALAKGEYSNKKITTVVDLKLAKSM 231

```

QY	227	MEQ	225
	:	:	:
Db	232	IEK	234

RESULT: 7
US-09 328-352-5107
Sec 5107, Application US/09328352

GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Que. / Match      8.7%; Score 104; DB 4; Length 461;
Best Local Similarity 21.7%; Pred. NO. 0.0017;
Matches 64; Conservative 38; Mismatches 75; Indels 118; Gaps 16

```

```

QY 1 KNPITIGCAWVAAGVGRMOADRPKOYLPFAGTVAIEHTLTRELSADEFQKVAIVSE 60
Dh 8 M8TTV-----TILAAKGTFRMSQLPKYLOPAGRPBIGHYI-----KTAQOLAE 53
QY 61 DYPWPELSIAKHPIITTAPOGKERADSVLSALKALEDIASEN-DWLVHDAARPCLTGSD 119
Dh 54 -----NITTYG--HOGDAH-----KTFPAQENIQWV--EQAEOLGTGHA 89
QY 120 IHQIDTLKANDPVG---GILASSHTLK-----HUNDGT----- 151
Dh 90 VQMTLPFVLPKGOISLILYGVDFVLRQTTLEQLI EASNKGTIGMTITLVHNPFGYGRIVRQ 149
QY 152 -----ITATIDRK-----HYMRALTPFOWKXGMLADALQRTGQ 184
Dh 150 DQKIQIAYBHNDATAGQIOEINTGTGYCVSNAKLHEW---LPKL-----SNENNOG 198
QY 185 NPAVTD-----BASALFELGHKPKI---VEGRPDNIKITRPEBLALAQYMEQQA 231
Dh 199 EYVLVDIVAMVAADGLBIASIQPELAEVEGVANDRLQLA-----ALEREFGQQA 248

RESULT: 8
US-O9_-98-452A-487
; Seq nce 487, Application US/09198452A
; Patent No. 6559294
; GENPAT. INFORMATION:
; APLICANT: Gillefais, R.

```

;; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
;; TITLE OF INVENTION: and treatment of infection
;; FILE REFERENCE: 9710-003-999
;; CURRENT APPLICATION NUMBER: US/09/198,452A
;; CURRENT FILING DATE: 1998-11-24
;; NUMBER OF SEQ ID NOS: 6849
;; SEQ ID NO 487
;; LENGTH: 357
;; TYPE: PRT
;; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-487

Query Match 7.8%; Score 93; DB 4; Length 357;
Best Local Similarity 26.0%; Pred. No. 0.023;
Matches 51; Conservative 30; Mismatches 79; Indels 36; Gaps 11;

QY 11 VPAGVGKRMQ-ADREKQY-----FLAGKVIHTLRLLESDAFQKVAVAISVEDPY 63
DB 92 LPPEAVSQILELSDGYSYLVNVRCDLSGETVQORL-LINADSGRSTVVIS-----B 145
QY 64 WPELSIAKHPDI-----ITAPGKERADSVLSALKALEDIASENDVVLVHDAPPC 114
DB 146 LPE-----GHPDIRNQLASERIFVSRERADAVASCKVAFPDHILPWSSHIAVAE 201
QY 115 LTGSDIHQITLKNQDPVGGILALSHDTLKHVDGDTITATIDKHYWRALTPQMEKYG 174
DB 202 IREKQOTMOGSLTEBOIGALIC-NTVSTKRL-AFALDAVL-KQSWRRRNDPLFAVE- 257
QY 175 LRDALQRTGNPAVTD 190
DB 258 -RALE-----ASVTD 267

RESULT 9
US-09-107-532A-6010
; Sequence 6010, Application US/09107532A
; Patent No. 6583275
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSES: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,596
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denise
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: CTC-012
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6010:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Enterococcus faecium
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (B) LOCATION 1...393
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6010:
US-C-107-532A-6010

QY Match 7.7%; Score 92; DB 4; Length 393;
Best Local Similarity 21.6%; Pred. No. 0.035;
Matches 53; Conservative 30; Mismatches 72; Indels 90; Gaps 9;

QY 8 WAPVAGVGKRMQADSPKQYLPAGKVIHTLRLLESDAFQKVAVAISVEDPYWPEL 67
DB 17 YAIILAKGRTMSKLYKYLHPVSGKPMTEHILNRVSEI-----SALKALEDIASENDVVLVHD 56
QY 68 SIARPDITAPG-GKERADSVL-----SALKALEDIASENDVVLVHD 109
DB 57 ---KPEVITIVGHAGQVKAQLGERSKVALQAEQGTGAVALQASFLSKGKGTIVIS 113
QY 110 AARPLTGSQIHLDITLKNQDPVGGILALSHDTLKHVDGDTITATIDKHYWRALTPQMEKYG 160
DB 114 GDTPLT-----TEFLN-----LREYHOGKNAATILTRQANPPTGGRITRDH 158
QY 161 V-----WRALTPQ-----MFYKGLRDALQRTGNPA-----VTDEASAL 195
DB 159 IGIVEKIVEQDAPFEALVQGIINTGYCFQNEALFALSKVGTNNAGGYLTDILIEL 218
QY 196 ELGGH 200
DB 219 KEKGH 223

RESULT 10
US-0-311-731A-34
; Sequence 34, Application US/08311731A
; Patent No. 6583266
GENERAL INFORMATION:

APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSES: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:

NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:

LENGTH: 285 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: MYCOBACTERIUM LEPRAE
 US-08-311-731A-34

Query Match 7.5%; Score 89.5; DB 4; Length 285;
 Best Local Similarity 18.9%; Pred. No. 0.042;

Matches 55; Conservative 43; Mismatches 102; Indels 91; Gaps 9;

QY 3 PTIQCAVVP-----AAGVCKMQADRPQYIPLAGKTYIETTLRLLESAFQ 51
 DB 10 PTLMAAILADTSNKNVCCGAGGKTTTAAAMALQAAEYGRIVV-----VLTIDPAK 63
 QY 52 KVAVAISVED-----PYMPELSIAKH-----PDIITAPGKERADSVLS--- 90
 DB 64 RLAAQLGVNDLGNTPQRFVPLAEVPGELHAMMLDMRRTFDEMNVVQYSGPERAAQAILNSEF 123
 QY 91 -----ALKALEDIASENDVLY---HDAAPCLTGSIDIHQIDTLKNDPV 132
 DB 124 YQTVATSLAGTQEVYAMERLQGLLSQDRMDLVVDTPPSRNALDFLDAPKRLGNFMNSRL 183
 QY 133 GGLIALSHDILKHYDGDITATTIDRKHWRALTPQKFKYGMALDALQRTGNDPAVTDEA 192
 DB 184 GRLILTPG-----RGIGRLVTGM---GLAMRALSTVIGSQMLADAA 222
 QY 193 SALELLGHRKPKIVEGRPDN-----IKITRPEDIALAQFYMEQ 229
 DB 223 TFVQSLDATTGSGFRKADRTYALTKQRTQFVVVSAAPDRLREASPFVDR 273

RESULT 11

US-08-311-731A-36
 Sequence 36, Application US/08311731A
 Patent No. 6583266

GENERAL INFORMATION:
 APPLICANT: SMITH, DOUGLAS

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
 NUMBER OF SEQUENCES: 411
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 STREET: 600 ATLANTIC AVENUE
 CITY: BOSTON
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/311,731A
 FILING DATE:

CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: GATES, EDWARD R.
 REGISTRATION NUMBER: 31,616

REFERENCE/DOCKET NUMBER: C0044/7125
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/720-3500
 TELEFAX: 617/720-2441

INFORMATION FOR SEQ ID NO: 36:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 285 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: MYCOBACTERIUM LEPRAE
 US-08-311-731A-36

Query Match 7.5%; Score 89.5; DB 4; Length 285;
 Best Local Similarity 18.9%; Pred. No. 0.042;

Matches 55; Conservative 43; Mismatches 102; Indels 91; Gaps 9;

QY 3 PTIQCAVVP-----AAGVCKMQADRPQYIPLAGKTYIETTLRLLESAFQ 51
 DB 10 PTLMAAILADTSNKNVCCGAGGKTTTAAAMALQAAEYGRIVV-----VLTIDPAK 63
 QY 52 KVAVAISVED-----PYMPELSIAKH-----PDIITAPGKERADSVLS--- 90
 DB 64 RLAAQLGVNDLGNTPQRFVPLAEVPGELHAMMLDMRRTFDEMNVVQYSGPERAAQAILNSEF 123
 QY 91 -----ALKALEDIASENDVLY---HDAAPCLTGSIDIHQIDTLKNDPV 132
 DB 124 YQTVATSLAGTQEVYAMERLQGLLSQDRMDLVVDTPPSRNALDFLDAPKRLGNFMNSRL 183
 QY 133 GGLIALSHDILKHYDGDITATTIDRKHWRALTPQKFKYGMALDALQRTGNDPAVTDEA 192
 DB 184 GRLILTPG-----RGIGRLVTGM---GLAMRALSTVIGSQMLADAA 222
 QY 193 SALELLGHRKPKIVEGRPDN-----IKITRPEDIALAQFYMEQ 229
 DB 223 TFVQSLDATTGSGFRKADRTYALTKQRTQFVVVSAAPDRLREASPFVDR 273

RESULT 12

US-08-936-165A-528
 Sequence 528, Application US/08936165A
 Patent No. 6348582

GENERAL INFORMATION:
 APPLICANT: Black, Michael

APPLICANT: Burnham, Martin
 APPLICANT: Hodgson, John
 APPLICANT: Knowles, David
 APPLICANT: Lometto, Michael
 APPLICANT: Nicholas, Richard
 APPLICANT: Pratt, Julie
 APPLICANT: Reichberg, Richard
 APPLICANT: Rosenberg, Martin
 APPLICANT: Ward, Judith
 TITLE OF INVENTION: No. 6348582e1 Prokaryotic Polynucleotides,
 TITLE OF INVENTION: Polypeptides and their uses
 NUMBER OF SEQUENCES: 534
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/936,165A
 FILING DATE: 24-SEP-1997

CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimm, Edward R.
 REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50549
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 60/027,032
 TELEFAX: 24-SEP-1996

TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 528:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 154 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-08-936-165A-528

Query Match 7.4%; Score 89; DB 4; Length 154;
 Best Local Similarity 23.9%; Pred. No. 0.13;
 Matches 26; Conservative 27; Mismatches 52; Indels 6; Gaps 4;

QY 80 GGERADSVLSAKLEDI--ASENDVYVHDAAPCLTGSQIHQITLKNDPVGGI-L 136
 DB 38 GGDRETTIMNIDHIRVNGIINDVITVDVAPPLQRIKENIEVAXX--YGAVDY 95

QY 137 ALSHDTL-KHYDGTITATIDRKAVRALTPQKFKYGMALDQRTGNP 186
 DB 96 VIEADITVMSKDKXNHSIPRNMETGQGTQSFNKLDDSYALSSSP 146

RESULT 13
 US-09-134-001C-3574
 Sequence 3574, Application US/09134001C
 Patent No. 6380370
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 3574
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3574

Query Match 7.4%; Score 89; DB 4; Length 531;
 Best Local Similarity 22.9%; Pred. No. 0.13;
 Matches 49; Conservative 37; Mismatches 72; Indels 56; Gaps 11;

QY 29 LPLAGTVEHTITRLLESDFQKAVAISSVEDPYWPELSAKHPDITTAQGERA--- 85
 DB 178 LPAABDPRVKQELIEDVIGID--QEDVYLASAKSNIGIEIEKIVDVVAPDGGPEAPLK 235
 QY 86 ----DS-----VLSAKLEDIASENDVYVHDAAPCLTGSQI-----HIQ 123
 DB 236 ALTFDESDPYRQVISIRIIDGVYAGRI-----KMAAGCKEFVTEVGIMPKPLP 289
 QY 124 IDTLKNDPVGGIALLSHDTLKAVD---GDTITAT--IDRK-HYVETALTPQNF----- 170
 DB 290 VEEITVGDVGYIIA-----SIKVVDSRGVDITITLAERPADPKLOGYKXKMPVFCGLP 344
 QY 171 ----KYGMALDQRTGNPAVTD--EASALELIG 199
 DB 345 IDKQYNDLREALEKQLNDASIEFFPSSQALG 378

RESULT 14
 US-08-971-782-2
 Sequence 2, Application US/08971782
 Patent No. 6043071
 GENERAL INFORMATION:

APPLICANT: Wallis, Nicola G.
 APPLICANT: Shilling, Lisa K.
 APPLICANT: Jaworski, Deborah D.
 APPLICANT: Wang, Min
 APPLICANT: Mooney, Jeffrey L.
 APPLICANT: Debonuck, Christine M.
 APPLICANT: Zhong, Yi Yi
 TITLE OF INVENTION: No. 6043071el GIMU
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Dechert Price & Rhoads
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
 CITY: Philadelphia
 STATE: PA
 COUNTRY: US
 ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/971,782
 FILING DATE:

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/050,996
 FILING DATE: 26-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Dickinson, Todd O
 REGISTRATION NUMBER: 28,354
 REFERENCE/DOCKET NUMBER: GM10024
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-994-2252
 TELEFAX: 215-994-2222

TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-971-782-2

Query Match 7.4%; Score 88.5; DB 3; Length 459;
 Best Local Similarity 23.9%; Pred. No. 0.12;
 Matches 48; Conservative 32; Mismatches 74; Indels 47; Gaps 9;

QY 8 MAVVPAAGYGMQADRPQYPLPLAGTVEHTITRLLESDFQKAVAISSVEDPYWPEL 67
 DB 4 FATLLAGGTNNKSLPVLKRVAGISLKH-----VRSVG-AIGPE----- 46
 QY 68 SIKHPDITTAQGERADSVLSAKLEDIASENDVYVHDA---APCLTGSQIHQI 124
 DB 47 -----KTVTVGHR--AEIVAEVLAGQTFVYQSEOGTGAHVMTPEPILEGSGHTLV 98
 QY 125 DTLKNDPVGGIALLSHDTLKAVDGTITATIDRKAVRALTPQV---FKYGMAL-----R 176
 DB 99 -----IAGDPPLTIGESLKNL-----IDFHNHNVATILVLAETDNPFGVGRIVRNDA 147
 QY 177 DALQRTGNPAVTDASALEL 197
 DB 148 EVLRVVEQDADTFEKOIKEI 168

RESULT 15
 US-09-309-026-2
 Sequence 2, Application US/09309026
 Patent No. 6204042
 GENERAL INFORMATION:
 APPLICANT: Wallis, Nicola G.
 APPLICANT: Shilling, Lisa K.
 APPLICANT: Jaworski, Deborah D.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 (Search time 42.325 Seconds
(without alignments)
3044.503 Million cell updates/sec

Title: US-09-941-947a-6
Perfect score: 3216
Sequence: 1 MKLTDYPLKNIHFPADIR.....LSLVGLSKGLATIEQPCA 620

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3216	100.0	620	10	US-09-934-903-2	Sequence 2, Appl1
2	3216	100.0	620	10	US-09-934-868-62	Sequence 62, Appl1
3	3216	100.0	620	11	US-09-941-947a-6	Sequence 6, Appl1
4	2068	64.7	628	12	US-10-369-493-13998	Sequence 13998, A
5	2068	64.4	627	12	US-10-381-779-15	Sequence 15, Appl1
6	2014.5	62.6	620	9	US-09-815-242-14083	Sequence 14083, A
7	2004.5	62.3	619	12	US-10-369-493-181	Sequence 481, App
8	2003.5	62.3	620	12	US-10-369-493-21174	Sequence 21174, A
9	2000.5	62.2	620	9	US-09-815-242-10068	Sequence 10068, A
10	2000.5	62.2	620	12	US-10-381-779-31	Sequence 31, Appl1
11	1990.5	62.0	620	12	US-10-369-493-731	Sequence 731, App
12	1992.5	62.0	623	12	US-10-369-493-15911	Sequence 15911, A
13	1992.5	62.0	623	12	US-10-369-493-16281	Sequence 16281, A
14	1992.5	62.0	653	12	US-10-369-493-15540	Sequence 15540, A
15	1990.5	61.9	619	12	US-10-369-493-7428	Sequence 7428, Ap

16	1985.5	61.7	616	12	US-10-369-493-4669	Sequence 4669, Ap
17	1971	61.3	615	12	US-10-369-493-9272	Sequence 9272, Ap
18	1968.5	61.2	618	12	US-10-369-493-8507	Sequence 8507, Ap
19	1967	61.2	613	12	US-10-369-493-9367	Sequence 9367, Ap
20	1951	60.7	670	12	US-10-369-493-17644	Sequence 17644, A
21	1937.5	60.2	625	12	US-10-381-779-33	Sequence 33, Appl1
22	1776	55.2	637	12	US-10-381-779-32	Sequence 32, Appl1
23	1623	50.5	624	12	US-10-369-493-10188	Sequence 10188, A
24	1620.5	50.4	628	12	US-10-381-779-18	Sequence 18, Appl1
25	1620.5	50.4	628	12	US-10-369-493-61	Sequence 61, Appl1
26	1555	48.4	628	12	US-10-369-493-20596	Sequence 20596, A
27	1543.5	48.0	637	12	US-10-369-493-12290	Sequence 12290, A
28	1532.5	47.7	629	12	US-10-381-779-3	Sequence 3, Appl1
29	1530.5	47.6	629	12	US-10-369-493-10848	Sequence 10848, A
30	1523.5	47.4	640	12	US-10-369-493-16942	Sequence 16942, A
31	1503.5	46.8	654	12	US-10-369-493-7893	Sequence 7893, Ap
32	1499.5	46.6	630	12	US-10-369-493-14775	Sequence 14775, A
33	1499.5	46.6	634	12	US-10-369-493-11753	Sequence 11753, A
34	1499.5	46.6	635	12	US-10-369-493-14237	Sequence 14237, A
35	1499.5	46.6	635	12	US-10-369-493-14928	Sequence 14928, A
36	1491.5	46.4	648	12	US-10-381-779-27	Sequence 27, Appl1
37	1485.5	46.2	720	12	US-10-427-590-26	Sequence 26, Appl1
38	1482	46.1	625	12	US-10-369-493-18033	Sequence 18033, A
39	1468	45.6	614	12	US-10-369-493-10317	Sequence 10317, A
40	1459	45.4	629	12	US-10-369-493-17357	Sequence 17357, A
41	1444	44.9	635	12	US-10-369-493-20015	Sequence 20015, A
42	1444	44.9	671	12	US-10-259-194A-414	Sequence 414, Appl1
43	1442	44.8	640	12	US-10-381-779-29	Sequence 29, Appl1
44	1442	44.8	640	12	US-10-369-493-2645	Sequence 2645, Ap
45	1439	44.7	636	12	US-10-381-779-28	Sequence 28, Appl1

ALIGNMENTS

RESULT 1
US-09-934-903-2
Sequence 2, Application US/09934903
Patent No. US20020102690A1
GENERAL INFORMATION:
APPLICANT: Kofias, Matheos
APPLICANT: Odou, J. Martin
APPLICANT: Schenzle, Andreas J.
APPLICANT: Temd, Jean-Francois
APPLICANT: Rouviere, Pierre
APPLICANT: Picataggio, Stephen
APPLICANT: Cheng, Olong
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
FILE REFERENCE: CL1646 US NA
CURRENT APPLICATION NUMBER: US/09/934,903
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: September 1, 2001
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 620
TYPE: PRT
ORGANISM: Methylobionas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by ORF1
US-09-934-903-2

Query Match 100.0% Score 3216; DB 10; Length 620;
Best Local Similarity 100.0% Pred. No. 1.1e-298; Indels 0; Gaps 0;
Matches 620; Conservative 0; Mismatches 0;
QY 1 MKLTDYPLKNIHFPADIRALSKQDLADAVGVYTHVVISGGHFAAGLTV 60
DB 1 MKLTDYPLKNIHFPADIRALSKQDLADAVGVYTHVVISGGHFAAGLTV 60
QY 61 ALHYVNTVVDQLVGHQAFPHILVGRKRMFTITLGGVSAFPAADSEVYAPCG 120

DB 61 ALHYFNTPVQDLVWDVGHQAVPEKILTGKREMPITRIITIGSVAPAPABESYDAFVG 120
QY 121 HSSSTISAAIGMAIASOLRGEDKKVVAIIDGSIITGMAVEANNAHAGDVANLLVILND 180
DB 121 HSSSTISAAIGMAIASOLRGEDKKVVAIIDGSIITGMAVEANNAHAGDVANLLVILND 180
QY 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKALAMPBWELARTBEHVGMIVPGT 240
DB 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKALAMPBWELARTBEHVGMIVPGT 240
QY 241 LFEELGFRVFGPIGDHVENMLVSTLEMLKDLTGPFVFLHVTYKKGKGAAPAKDPLAYHGV 300
DB 241 LFEELGFRVFGPIGDHVENMLVSTLEMLKDLTGPFVFLHVTYKKGKGAAPAKDPLAYHGV 300
QY 301 PAFPTKDFLPKAPSPHPTTYTEVFGWMLCDMAODERLLGITPAMRBSGLVEFSQKFP 360
DB 301 PAFPTKDFLPKAPSPHPTTYTEVFGWMLCDMAODERLLGITPAMRBSGLVEFSQKFP 360
QY 361 NRYEDVAIAEQHAVALTAAGACQAKPVVAIYSTFLORGVDOLIHVALONTLMFLALDR 420
DB 361 NRYEDVAIAEQHAVALTAAGACQAKPVVAIYSTFLORGVDOLIHVALONTLMFLALDR 420
QY 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFOHGGASVRYPRGKG 480
DB 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFOHGGASVRYPRGKG 480
QY 481 PGAAIDPPLTALBEGKAEVRHGGSRILILANGSVTPAVAGKQLGATVNNKPFVPEFDQ 540
DB 481 PGAAIDPPLTALBEGKAEVRHGGSRILILANGSVTPAVAGKQLGATVNNKPFVPEFDQ 540
QY 541 ALVLELATRHVFTVEENVTAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600
DB 541 ALVLELATRHVFTVEENVTAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600
QY 601 LSLVGLDSKGLATITEQFCA 620
DB 601 LSLVGLDSKGLATITEQFCA 620

RESULT 2
US-09-934-868-62
Sequence 62, Application US/09934868
Patent No. US20020137130A1
GENERAL INFORMATION:
APPLICANT: Koffas, Mattheos
APPLICANT: Odom, James M.
APPLICANT: Schenazle, Andreas J.
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: CL1596 US NA
CURRENT APPLICATION NUMBER: US/09/934,868
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft Office 97
SEQ ID NO 62
LENGTH: 620
TYPE: prt
ORGANISM: Methylobionas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by DXS
US-09-934-868-62

Query Match 100.0%; Score 3216; DB 10; Length 620;
Best Local Similarity 100.0%; Pred. No. 1,1e-298;
Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLTTDYPLLNKNIHTPADIRALSKDQLQDLADBVGVYTHTVSISGGHFAAGCTVETLV 60
DB 1 MKLTTDYPLLNKNIHTPADIRALSKDQLQDLADBVGVYTHTVSISGGHFAAGCTVETLV 60
QY 61 ALHYFNTPVQDLVWDVGHQAVPEKILTGKREMPITRIITIGSVAPAPABESYDAFVG 120

DB 61 ALHYFNTPVQDLVWDVGHQAVPEKILTGKREMPITRIITIGSVAPAPABESYDAFVG 120
QY 121 HSSSTISAAIGMAIASOLRGEDKKVVAIIDGSIITGMAVEANNAHAGDVANLLVILND 180
DB 121 HSSSTISAAIGMAIASOLRGEDKKVVAIIDGSIITGMAVEANNAHAGDVANLLVILND 180
QY 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKALAMPBWELARTBEHVGMIVPGT 240
DB 181 DMSISPPVGMANNLITVLSKFPYSSVRESKKALAMPBWELARTBEHVGMIVPGT 240
QY 241 LFEELGFRVFGPIGDHVENMLVSTLEMLKDLTGPFVFLHVTYKKGKGAAPAKDPLAYHGV 300
DB 241 LFEELGFRVFGPIGDHVENMLVSTLEMLKDLTGPFVFLHVTYKKGKGAAPAKDPLAYHGV 300
QY 301 PAFPTKDFLPKAPSPHPTTYTEVFGWMLCDMAODERLLGITPAMRBSGLVEFSQKFP 360
DB 301 PAFPTKDFLPKAPSPHPTTYTEVFGWMLCDMAODERLLGITPAMRBSGLVEFSQKFP 360
QY 361 NRYEDVAIAEQHAVALTAAGACQAKPVVAIYSTFLORGVDOLIHVALONTLMFLALDR 420
DB 361 NRYEDVAIAEQHAVALTAAGACQAKPVVAIYSTFLORGVDOLIHVALONTLMFLALDR 420
QY 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFOHGGASVRYPRGKG 480
DB 421 AGLVGPDGPTHAGAFDYSYKRCIPNMLIMAPADENECROMLTGFOHGGASVRYPRGKG 480
QY 481 PGAAIDPPLTALBEGKAEVRHGGSRILILANGSVTPAVAGKQLGATVNNKPFVPEFDQ 540
DB 481 PGAAIDPPLTALBEGKAEVRHGGSRILILANGSVTPAVAGKQLGATVNNKPFVPEFDQ 540
QY 541 ALVLELATRHVFTVEENVTAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600
DB 541 ALVLELATRHVFTVEENVTAGAGSAINTFLOAQKVLMPVCNIGLPRFVQGSREEL 600
QY 601 LSLVGLDSKGLATITEQFCA 620
DB 601 LSLVGLDSKGLATITEQFCA 620

RESULT 3
US-09-941-947a-6
Sequence 6, Application US/09941947A
Publication No. US20030003528A1
GENERAL INFORMATION:
APPLICANT: Brzostowicz, Patricia C.
APPLICANT: Cheng, Qiong
APPLICANT: DiCosimo, Deana J.
APPLICANT: Koffas, Mattheos
APPLICANT: Miller, Edward S. Jr.
APPLICANT: Odom, J. Martin
APPLICANT: Picataggio, Steve
APPLICANT: Rouviere, Pierre E.
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
FILE REFERENCE: CL1903 US NA
CURRENT APPLICATION NUMBER: US/09/941,947A
PRIOR FILING DATE: 2001-09-01
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 620
TYPE: prt
ORGANISM: Methylobionas 16a
US-09-941-947a-6

Query Match 100.0%; Score 3216; DB 11; Length 620;
Best Local Similarity 100.0%; Pred. No. 1,1e-298;
Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKLTTDYPPLKNIHTPADIRALSKDQLOQADREVGYLTHTVSISGGHFAAGLTVELTV 60
Db 1 MKLTTDYPPLKNIHTPADIRALSKDQLOQADREVGYLTHTVSISGGHFAAGLTVELTV 60
QY 61 ALHYFNTPTVDQVMDVGHQAVPHKILNCRKEMPTITRLGGVSAFPARDESEYDAPGVG 120
Db 61 ALHYFNTPTVDQVMDVGHQAVPHKILNCRKEMPTITRLGGVSAFPARDESEYDAPGVG 120
QY 121 HSTSTISALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDN 180
Db 121 HSTSTISALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDN 180
QY 121 HSTSTISALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDN 180
Db 121 HSTSTISALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDN 180
QY 181 DMSISPPVGMANNVLTLYLSKFTSSVRESKALAKMPSWELARKTEEHVKMIVPGT 240
Db 181 DMSISPPVGMANNVLTLYLSKFTSSVRESKALAKMPSWELARKTEEHVKMIVPGT 240
QY 241 LFEELGPNYFGPIIDGHVEMVSTLENLKDLPYPLVHTVTKKGVAAPAKDPLAHGV 300
Db 241 LFEELGPNYFGPIIDGHVEMVSTLENLKDLPYPLVHTVTKKGVAAPAKDPLAHGV 300
QY 301 PAFDPTKDFLPRAADSPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 360
Db 301 PAFDPTKDFLPRAADSPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 360
QY 361 NRYFVAIAEQAHTVLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 420
Db 361 NRYFVAIAEQAHTVLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 420
QY 421 AGVGPDPPTAGADPVSYWRCIPNMLIMAPADNECROMLTGPOHGHASVRYRSGKGA 480
Db 421 AGVGPDPPTAGADPVSYWRCIPNMLIMAPADNECROMLTGPOHGHASVRYRSGKGA 480
QY 481 PGAAIDPPLTALIEIKAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVVMREVPKFDQ 540
Db 481 PGAAIDPPLTALIEIKAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVVMREVPKFDQ 540
QY 541 ALVLELARTHVFTVEENVAAGGSAINTFLQAKVLMPCVNICIGLPPRVGSGSEEL 600
Db 541 ALVLELARTHVFTVEENVAAGGSAINTFLQAKVLMPCVNICIGLPPRVGSGSEEL 600
QY 601 LSLVGLDSKGLATIEQFCA 620
Db 601 LSLVGLDSKGLATIEQFCA 620

RESULT 4
US-10-369-493-13998
; Sequence 13998, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-19(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13998
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13998

Query Match 64.7%; Score 2080; DB 12; Length 628;
Best Local Similarity 63.2%; Pred. No. 7e-190;
Matches 387; Conservative 101; Mismatches 120; Indels 4; Gaps 2;

```

```

QY 8 PULKNIHTPADIRALSKDQLOQADREVGYLTHTVSISGGHFAAGLTVELTV 67
Db 16 PULDRANTPDGRRGRHALETALDELDELITYTGQCGHGAAGLTVELTV 75
QY 68 TPVDQVMDVGHQAVPHKILNCRKEMPTITRLGGVSAFPARDESEYDAPGVGHS 127
Db 76 TPDDRLVMDVGHQAVPHKILNCRKEMPTITRLGGVSAFPARDESEYDAPGVGHS 135
QY 128 AALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDNDSISPP 187
Db 136 AALGMAISQLEGEDEKRWVAIIGDSITGMAVEAMNAGVANNLITLINDNDSISPP 195
QY 188 VGMANNVLTLYLSKFTSSVRESKALAKMPSWELARKTEEHVKMIVGGLPEELG 247
Db 196 VGLSNVYAKLISSTVYASMRGSKVLSRLPGAEIARREEYAKMIVGGLPEELG 255
QY 248 NRYFVAIAEQAHTVLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 307
Db 256 NRYFVAIAEQAHTVLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 315
QY 308 DFLPRAAP--SHPHYTYEVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 365
Db 316 --APAAAPKAGGPKYSGVFGNMLCDMAADERLLGITPAMEGSGLVEFSQKFP 373
QY 366 VALAQHVAITLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 425
Db 374 VALAQHVAITLAAGQACQAKPVVAIYSTFLQGVQDLIHVALQNLDMLEFALDR 433
QY 426 PGDPPTAGADPVSYWRCIPNMLIMAPADNECROMLTGPOHGHASVRYRSGKGA 485
Db 434 EDPPTAGADPVSYWRCIPNMLIMAPADNECROMLTGPOHGHASVRYRSGKGA 493
QY 486 DPTLALBIGAEVRHGSRIALIAWGSMTTPAVEAGKOLGATVVMREVPKFDQ 545
Db 494 EDPPTAGADPVSYWRCIPNMLIMAPADNECROMLTGPOHGHASVRYRSGKGA 553
QY 546 LARTHDVPTVEENVAAGGSAINTFLQAKVLMPCVNICIGLPPRVGSGSEEL 605
Db 554 LANSDDLVTLEENVAAGGSAINTFLQAKVLMPCVNICIGLPPRVGSGSEEL 613
QY 606 LDSKGLATIEQ 617
Db 614 LDEVGIEAAVRE 625

RESULT 5
US-10-381-779-15
; Sequence 15, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jessem, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12304/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-381-779-15

Query Match 64.4%; Score 2069.5; DB 12; Length 627;
Best Local Similarity 63.4%; Pred. No. 7.1e-189;
Matches 388; Conservative 100; Mismatches 117; Indels 7; Gaps 2;

```

QY 8 P L K N I H T P A D I R A L S K D O L Q O L A D E V R G Y L T H T V S I S G G H P A A G I G T V E L T A L A Y V N 67
 DB 16 P L L R A S S P A E L R L G E D L E T L A D E L R Q Y L L Y V G O T G H P G A G L V A L T A L H Y V D 75
 QY 68 T P V D O L V M D V G H Q A V P H K I L T G K E R M P T I R T L G V A P A D E S E Y D A F G V G S S T S I S 127
 DB 76 T P D R L V M D V G H Q A V P H K I L T G K E R M P T I R T L G V A P A D E S E Y D A F G V G S S T S I S 135
 QY 128 A A L G M A I A S Q L R G E D K K M A L I G D S I T G M A Y A M M A D V A N A L L V I L I N D M S I S P 187
 DB 136 A A L G M A I A R L Q G E R K S V A I G D A L T A G M A F E A L M A S E V A D M V I I N D M S I S H N 195
 QY 188 V G A M N N Y L T K V L S K F Y S S V R E S K A L A K M P S W E L A R K T E E H V K M I V P G T L F E S L G F 247
 DB 196 V G G A S N I A K L I S R T S S M A R E S K R U L S R L P G M E L A R K T E E H V A K M L V P G T L F E S L G M 255
 QY 248 N Y F G P I D G H D V E M L V S T L E N I K D I G P V L H V V T K K G Y A P A K O P L A Y H G V P A D P T K 307
 DB 256 N Y I G P I D G H D L P T L V A T L R M R D M K G P O F L H V V T K K G Y A P A L D P I G Y H A I - - - - T K 310
 QY 308 D E L P R A A R - - S H P T Y T E V P G R W L C D M A A O E R L I G T P M A R E S G L V E R S O K P P N R Y E D 365
 DB 311 L E A R S A K K T G G P T S S V F Q M L C D M A A Q D A R L I G T P M A R E S G L V A E S E R Y P E R Y E D 370
 QY 366 V A I A Q H A V T L A A G A C Q G A R P V A I Y S T F L O R G Y D O L I H V A L Q N L M L F A L D R A G L V G 425
 DB 371 V A I A Q H A V T L A A G A C G M K P V A I Y S T F L O R A Y D O L I H V A Q H D V F A L D R A G L V G 430
 QY 426 P D G P T H A G A P Y S Y R K C I P N M L I A P A D E N E C R O M L T T G R H G P A S R V R P R G K G A A I 485
 DB 431 E D G P T H A S F D I S Y R C I P G M L V M T P S D E L R K L T T G V L P D P A A V R P R S G G P E R P I 490
 QY 486 D P T L A L E I G K A E V H G S R I A I L A M G S M V T P A V E A K O L G A T V V N R F V K P P D O A L V E 545
 DB 491 D P D L O P V E I G K V R R R G R V A L L F V G Q L A E A K M A B S I D A T V D R F V K P P I D E A L V R E 550
 QY 546 I A R T H D V F V T E E N V I A G A S A I N T F L Q A K V L M P V C N I G L P R F V G S R E B L L S V G 605
 DB 551 I A G S H E L L V T I B E N A V M G A S A V G E F L A S E L E V P L L Q L G P Y V Y E A H K P S E M A E C G 610
 QY 606 I D S K G I L A T I B O F C A 617
 DB 611 I D A G I E A K I K A M L A 622

RESULT 6
 US-09-815-242-14083
 ; Sequence 14083, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27

QY 7 Y P L K N I H T P A D I R A L S K D O L Q O L A D E V R G Y L T H T V S I S G G H P A A G I G T V E L T A L A Y V N 66
 DB 8 Y P T A L V D S T G E L L P R E S I P K C D L R R Y L D S V R S S G H P A A G I G T V E L T A L A Y V 67
 QY 67 N T P V D O L V M D V G H Q A V P H K I L T G K E R M P T I R T L G V A P A D E S E Y D A F G V G S S T S I 126
 DB 68 N T P D O L I M D V G H Q A V P H K I L T G R B D X I G T I R Q G L H P P M R G E S E Y D V L S G S S T S I 127
 QY 127 S A L I G M A I A S Q L R G E D K K M A L I G D S I T G M A Y A M M A D V A N A L L V I L I N D M S I S P 186
 DB 128 S A G I G I A V A A K E R D R T V C I D G A I T A G M A F E A N N H A G D I R P D M L V I I N D M S I S B 187
 QY 187 P V G A M N N Y L T K V L S K F Y S S V R E S K A L A K M P S W E L A R K T E E H V K M I V P G T L F E S L G 246
 DB 188 N V G A L N N L A L I S K L Y S I R E G K K V F S G V P I K I L L R K T E E H I K G V V P G T L F E S L G 247
 QY 247 F N Y F G P I D G H D V E M L V S T L E N I K D I G P V L H V V T K K G Y A P A K O P L A Y H G V P A D P T K 306
 DB 248 F N Y I G P V D G H D V M E L I S T A K M R D L K G P O F L I H T K K R G E P A E K D P I T F H A V P K E D P S 307
 QY 307 K D F L P K A P S H P Y T T E V P G R W L C D M A A O E R L I G T P M A R E S G L V E R S O K P P N R Y E D V 366
 DB 308 S G C P K S S - G G L P G S K I P G M L C S T A K S K M A L I T P A R E S G M E V S R K P P D R Y E D V 366
 QY 367 A I A B Q H A V T L A A G A C Q G A R P V A I Y S T F L O R G Y D O L I H V A L Q N L M L F A L D R A G L V P 426
 DB 367 A I A B Q H A V T P A A G I A I G G Y K P V A I Y S T F L O R A V D O L I H V A I Q L P V M P A I D R A G I V G A 426
 QY 427 D C P T H A G A P Y S Y R K C I P N M L I A P A D E N E C R O M L T T G P O H N - G P A S R V R P R G K G A A I 485
 DB 427 D C Q T H Q A P D L S Y R C I P D M V I M T P S D E N E C R O M L F T G Y H N D E P L A V R I P R G A O G V A L 486
 QY 486 D P T L A L E I G K A E V H G S R I A I L A M G S M V T P A V E A K O L G A T V V N R F V K P P D O A L V E 545
 DB 487 T P - L E K L I I G K G V Y R H E K A I L F G T L M P E A K V A B A L N A T I V D R F V K P P I D D T L I L E 545
 QY 546 I A R T H D V F V T E E N V I A G A S A I N T F L Q A K V L M P V C N I G L P R F V G S R E B L L S V G 605
 DB 546 M A A Q H D A L V T L E E N A I M G A S G V N E V L A R K R P V A L N I G L P D F I P Q G T O E S A R A E L G 605
 QY 606 I D S K G I L A T I B O F C A 620
 DB 606 I D A G I E A K I K A M L A 620

RESULT 7
 US-10-369-493-481
 ; Sequence 481, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; PLANTS WITH IMPROVED PROPERTIES

PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 14083
 ; LENGTH: 620
 ; TYPE: PRT
 ; ORGANISM: Salmonella typhi
 ; US-09-815-242-14083

Query Match 62.6%; Score 2014.5; DB 9; Length 620;
 Best Local Similarity 61.0%; Pred. No. 1.3e-183;
 Matches 375; Conservative 106; Mismatches 131; Indels 3; Gaps 3;

FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO 481
 LENGTH: 619
 TYPE: PR
 ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-481

Query Match 62.3%; Score 2004.5; DB 12, Length 619;
 Best Local Similarity 61.8%; Pred. No. 1.2e-182;
 Matches 378; Conservative 102; Mismatches 129; Indels 3; Gaps 3;

QY 7 YPLKNHPTPADIRALSKDQLQCLADEVRYLTHTVTSISGGHFAAGLGTVELTVALHYVF 66
 DB 8 YPTLALATPEBELRLIPKDTLPKCDLROPLANSVSSSGHFAAGLGTVELTVALHYVF 67
 QY 67 NTPVDQVWDVGHQAYPHKILTRKREMPRTITLGGVSAFPARDESEYDAFGVGHSTSI 126
 DB 68 KTFPDMVWDVGHQAYPHKILTRGRDRIINTIRKQGLHPFWRNESEYDTLCVGHSTSI 127
 QY 127 SALLGMAIALSQRGDKKCVAILIGDSITGGAAVYAMNHAGDVANLLVILINDMSTSP 186
 DB 128 SALLGMAIALAKEDKGRKTVCVIIGGALTAGAFAMNHAGDIDPDMVILINDMSTSP 187
 QY 187 PYGAMNNVLTLYLSKSFYSSVYESSKALAKMPVSVELARTBEHVKMIVPGLTFEELG 246
 DB 188 NVGALNNHLAHLSSKLTTLREGSKYFSGLPKIKLKTTEHLKGMVPGTLFEEIG 247
 QY 247 FNYFPGIDGHVEMVSTLLENKDI TGVPFLHVTTKKGKGYAPAKOPLAYHGVPAFPT 306
 DB 248 FNYFPGIDGHVEMVSTLLENKDI TGVPFLHVTTKKGKGYAPAKOPLAYHGVPAFPT 307
 QY 307 KDFLKAAPSHPPTTEVFGRLCPMAQDERLLGITPRAMEGSGLVESOKFNRFPDY 366
 DB 308 TGSLLPKSSDT- RPTTSKIFGEMCEBAHDKKMAITPAMEGSGMRFSEYEQFEDV 366
 QY 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALONLMDLPALDRAGLVGP 426
 DB 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALONLMDLPALDRAGLVGP 426
 QY 427 DGPTRGAFDYSVMRCINMLIMAPDENECROMLTGFGCH--HBPASVYRPRGSPAAI 485
 DB 427 DGPTRGAFDYSVMRCINMLIMAPDENECROMLTGFGCH--HBPASVYRPRGSPAAI 486
 QY 486 DPTLTALBIRGAVERHHSRIAILAMGSMVTPAVEAGKOLGATVVMRFPVPEPDALYLE 545
 DB 487 CP-LEILPFGKIVRQGEKIALINFGTLLENALQAAESINAVVDRFPVPEPDALYLE 545
 QY 546 LARTHDVPTVBEENVLAGAGSAINFLQAKVILMPYCNIGLPRFPBGSSREELSLV 605
 DB 546 LARTHDVPTVBEENVLAGAGSAINFLQAKVILMPYCNIGLPRFPBGSSREELSLV 605
 QY 606 LDSKGLIATIEQ 617
 DB 606 LDSKGLIATIEQ 617

RESULT 8
 US-10-369-493-21174
 Sequence 21174, Application US/10369493
 Publication No. US2003023675A1
 GENERAL INFORMATION:
 APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO 21174
 LENGTH: 620
 TYPE: PR
 ORGANISM: Xenorhabdus nematophilus
 US-10-369-493-21174

Query Match 62.3%; Score 2003.5; DB 12, Length 620;
 Best Local Similarity 61.1%; Pred. No. 1.5e-182;
 Matches 377; Conservative 102; Mismatches 131; Indels 7; Gaps 4;

QY 7 YPLKNHPTPADIRALSKDQLQCLADEVRYLTHTVTSISGGHFAAGLGTVELTVALHYVF 66
 DB 8 YPTLALATPEBELRLIPKDTLPKCDLROPLANSVSSSGHFAAGLGTVELTVALHYVF 67
 QY 67 NTPVDQVWDVGHQAYPHKILTRKREMPRTITLGGVSAFPARDESEYDAFGVGHSTSI 126
 DB 68 KTFPDMVWDVGHQAYPHKILTRGRDRIINTIRKQGLHPFWRNESEYDTLCVGHSTSI 127
 QY 127 SALLGMAIALSQRGDKKCVAILIGDSITGGAAVYAMNHAGDVANLLVILINDMSTSP 186
 DB 128 SALLGMAIALAKEDKGRKTVCVIIGGALTAGAFAMNHAGDIDPDMVILINDMSTSP 187
 QY 187 PYGAMNNVLTLYLSKSFYSSVYESSKALAKMPVSVELARTBEHVKMIVPGLTFEELG 246
 DB 188 NVGALNNHLAHLSSKLTTLREGSKYFSGLPKIKLKTTEHLKGMVPGTLFEEIG 247
 QY 247 FNYFPGIDGHVEMVSTLLENKDI TGVPFLHVTTKKGKGYAPAKOPLAYHGVPAFPT 306
 DB 248 FNYFPGIDGHVEMVSTLLENKDI TGVPFLHVTTKKGKGYAPAKOPLAYHGVPAFPT 307
 QY 307 KDFLKAAPSHPPTTEVFGRLCPMAQDERLLGITPRAMEGSGLVESOKFNRFPDY 366
 DB 308 TGSLLPKSSDT- RPTTSKIFGEMCEBAHDKKMAITPAMEGSGMRFSEYEQFEDV 366
 QY 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALONLMDLPALDRAGLVGP 426
 DB 367 ALAEQHAVTLAAGACQCAKPVVAIYSTFLQRYDQLIHDVALONLMDLPALDRAGLVGP 426
 QY 427 DGPTRGAFDYSVMRCINMLIMAPDENECROMLTGFGCH--HBPASVYRPRGSPAAI 485
 DB 427 DGPTRGAFDYSVMRCINMLIMAPDENECROMLTGFGCH--HBPASVYRPRGSPAAI 486
 QY 486 DPTLTALBIRGAVERHHSRIAILAMGSMVTPAVEAGKOLGATVVMRFPVPEPDALYLE 543
 DB 487 CP-LEILPFGKIVRQGEKIALINFGTLLENALQAAESINAVVDRFPVPEPDALYLE 543
 QY 544 LARTHDVPTVBEENVLAGAGSAINFLQAKVILMPYCNIGLPRFPBGSSREELSLV 603
 DB 544 LARTHDVPTVBEENVLAGAGSAINFLQAKVILMPYCNIGLPRFPBGSSREELSLV 603
 QY 604 VGLDSKGLIATIEQ 620
 DB 604 VGLDSKGLIATIEQ 620

RESULT 9
 US-09-815-242-10068
 Sequence 10068, Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10068
LENGTH: 620
TYPE: PR1
ORGANISM: Escherichia coli
US-09-815-242-10068

Query Match 62.2% Score 2000.5; DB 9; Length 620;
Best Local Similarity 60.2%; Pred. No. 2.8e-182;
Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;

7 YPLNKHITPADIRALSKDQQLADEVRGLTHVTSISGHPAAGTVEITVALHYF 66
8 YPTLALVSTOELRALPESLPCDELRRLDSVSSSHFASGLGTVALHYF 67
67 NTPVQOLVWDVGHQAYPKHILTGKREKAPRTIRTLGVSAPAPADESEYDAGVGHSTSI 126
68 NTPVQOLVWDVGHQAYPKHILTGKREKAPRTIRTLGVSAPAPADESEYDAGVGHSTSI 127
127 SATGMAIASQLRGEDEKKNVAILIGDSITGGMAYEAMNADVNANLITINDMSISP 186
128 SAGIGIAYAAEKEGKNNRTVCYIGGALTAGMAPEANMAGDIPDMVLINDMSISP 187
187 PYGAMNNYITKYLSSKPYSSVSEESKALAMPVWEELAKTTEHVKNMIVPGLFEEIG 246
188 NYGALNNHLAGLSGLYSLSREGKVPVSGVPPKELKRTTEHVKMVPVGLFEEIG 247
247 FNYGPIIDGHVEMVSTLENKDLTGPFVFLHVTVKKGAYAPAKDPLAHGVAPDPPT 306
248 FNYGPIIDGHVEMVSTLENKDLTGPFVFLHVTVKKGAYAPAKDPLAHGVAPDPPT 307
307 KOTLPKAPSPHPTTEVGRWLCDMAADDERLLGITPAMEGSGLVFESQKPFNRYFDV 366
308 SCGLPKSS-CGLPSYSKIFGDMLCETAAKONKLAITPAMEGSGLVFESQKPFNRYFDV 366
367 AIAEGHNAVTLAAGAACGAKPVVAIYSTFLQRYGQDILHDVALQNLDMFLALDRAGLVGP 426
367 AIAEGHNAVTLAAGAACGAKPVVAIYSTFLQRYGQDILHDVALQNLDMFLALDRAGLVGP 426
427 DQPTAGADVSVMACIPNMLIMAPADENECOMLTTSPQH-GRASVRYRGKGPAAI 485
427 DQPTAGADVSVMACIPNMLIMAPADENECOMLTTSPQH-GRASVRYRGKGPAAI 486
486 DPLTLALBICAEVHRHGSRIAILMAGSVNTPAVAGKQIATVNMRFVYRPPDQALVLE 545
487 TP-LKPLPIGKIVRREGKLAIFNGTLMPEAPAVASINATLVDMFVPLDBALILE 545
546 LARTHDVFTVEENYIAGGASAINFTLOAQKVLMPVCNIGLPDRFVQSGREELLSVIG 605
546 MASHBALVTVENNALMGAGGVNEMARHKPVVINTIGLPDFFIQGTQOEBWRAELG 605
606 LDSKGLIATIBQFCA 620

Db 606 LDAAGMEKIRAMLA 620
RESULT 10
US-10-381-779-31
Sequence 31, Application US/10381779
Publication No. US20030219798A1
GENERAL INFORMATION:
APPLICANT: Jensen, Ravi R
APPLICANT: Zidwick, Mary Jo
TITLE OF INVENTION: Isoprenoid Production
FILE REFERENCE: 12904/002051
CURRENT APPLICATION NUMBER: US/10/381,779
CURRENT FILING DATE: 2003-03-28
PRIOR APPLICATION NUMBER: PCT/US01/30328
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/236,580
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 190
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 620
TYPE: PR1
ORGANISM: Escherichia coli
US-10-381-779-31

Query Match 62.2% Score 2000.5; DB 12; Length 620;
Best Local Similarity 60.2%; Pred. No. 2.8e-182;
Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;

7 YPLNKHITPADIRALSKDQQLADEVRGLTHVTSISGHPAAGTVEITVALHYF 66
8 YPTLALVSTOELRALPESLPCDELRRLDSVSSSHFASGLGTVALHYF 67
67 NTPVQOLVWDVGHQAYPKHILTGKREKAPRTIRTLGVSAPAPADESEYDAGVGHSTSI 126
68 NTPVQOLVWDVGHQAYPKHILTGKREKAPRTIRTLGVSAPAPADESEYDAGVGHSTSI 127
127 SATGMAIASQLRGEDEKKNVAILIGDSITGGMAYEAMNADVNANLITINDMSISP 186
128 SAGIGIAYAAEKEGKNNRTVCYIGGALTAGMAPEANMAGDIPDMVLINDMSISP 187
187 PYGAMNNYITKYLSSKPYSSVSEESKALAMPVWEELAKTTEHVKNMIVPGLFEEIG 246
188 NYGALNNHLAGLSGLYSLSREGKVPVSGVPPKELKRTTEHVKMVPVGLFEEIG 247
247 FNYGPIIDGHVEMVSTLENKDLTGPFVFLHVTVKKGAYAPAKDPLAHGVAPDPPT 306
248 FNYGPIIDGHVEMVSTLENKDLTGPFVFLHVTVKKGAYAPAKDPLAHGVAPDPPT 307
307 KOTLPKAPSPHPTTEVGRWLCDMAADDERLLGITPAMEGSGLVFESQKPFNRYFDV 366
308 SCGLPKSS-CGLPSYSKIFGDMLCETAAKONKLAITPAMEGSGLVFESQKPFNRYFDV 366
367 AIAEGHNAVTLAAGAACGAKPVVAIYSTFLQRYGQDILHDVALQNLDMFLALDRAGLVGP 426
367 AIAEGHNAVTLAAGAACGAKPVVAIYSTFLQRYGQDILHDVALQNLDMFLALDRAGLVGP 426
427 DQPTAGADVSVMACIPNMLIMAPADENECOMLTTSPQH-GRASVRYRGKGPAAI 485
427 DQPTAGADVSVMACIPNMLIMAPADENECOMLTTSPQH-GRASVRYRGKGPAAI 486
486 DPLTLALBICAEVHRHGSRIAILMAGSVNTPAVAGKQIATVNMRFVYRPPDQALVLE 545
487 TP-LKPLPIGKIVRREGKLAIFNGTLMPEAPAVASINATLVDMFVPLDBALILE 545
546 LARTHDVFTVEENYIAGGASAINFTLOAQKVLMPVCNIGLPDRFVQSGREELLSVIG 605
546 MASHBALVTVENNALMGAGGVNEMARHKPVVINTIGLPDFFIQGTQOEBWRAELG 605
606 LDSKGLIATIBQFCA 620

Db 606 LBAAGWEAKIKAWILA 620

RESULT 11

US-10-369-493-731
 : Sequence 731, Application US/10369493
 : Publication No. US20030233675A1
 : GENERAL INFORMATION:
 : APPLICANT: Cao, Yongwei
 : APPLICANT: Hinkle, Gregory J.
 : APPLICANT: Slater, Steven C.
 : APPLICANT: Goldman, Barry S.
 : APPLICANT: Chen, Xianfeng
 : TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 : FILE REFERENCE: 38-10(52052)B
 : CURRENT APPLICATION NUMBER: US/10/369,493
 : PRIORITY FILING DATE: 2003-02-28
 : PRIOR APPLICATION NUMBER: US 60/360,039
 : NUMBER OF SEQ ID NOS: 47374
 : SEQ ID NO 731
 : LENGTH: 620
 : TYPE: PRT
 : ORGANISM: Escherichia coli
 : US-10-369-493-731

Query Match 62.2%; Score 2000.5; DB 12; Length 620;
 Best Local Similarity 60.2%; Pred. No. 2.8e-182;
 Matches 370; Conservative 114; Mismatches 128; Indels 3; Gaps 3;

QY 7 YPLLNKIHHPADIRALSKOQLQDLADVSGYLTHTVSGSGHFAAGLGVETVALHYV 66
 DB 8 YPLALVDSIOEIRLPKESLPCLCELRRLYDLSVRSKSGHFAAGLGVETVALHYV 67
 QY 67 NTPVDOLVMDVGHQAVPHKILTGKREMPFIRTLGVSAPFARDESSEYDAFGVGSSTSI 126
 DB 68 NTPFDOLVMDVGHQAVPHKILTGKREMPFIRTLGVSAPFARDESSEYDAFGVGSSTSI 127
 QY 127 SALLGMAIALSOLGEGEDKKNVAILIGDGSITGGMAVEANAGDVMANLVLINDMSISP 186
 DB 128 SAIGIGIAVAEKEGKRRVYCVIGDAITAGMAFEAMNAGDIRPDMVYLINDMSISB 187
 QY 187 PVGAMNNYLTQVLSKFSYSSVRESKKALAKMPSVWEILAKTEEHYKGVITVGTFFELG 246
 DB 188 NVGALNNHIAQLISGLYSLSREGGKVFSGVPPIBELKRTESHIKGVVFGTILFEEIG 247
 QY 247 FNYGPIIDGHDVEMVSTLENLKDILTGPFVLTHTVTKKGVAFAEKDPLAVGVPADPT 306
 DB 248 FNYIGVGDHVDVGLITTLKNMRDLKGQPLHIMTKKGVSFAEKDPTFAVVPKDPSS 307
 QY 307 KDLPLKAPSPHPTTYEVGKRLCDMAADDERLGLITPAMREGSLVFSQTFPRRYDV 366
 DB 308 SGCLPKRS-GLPSYSKIFGDMCEPAADNKLMAITPAMREGSGMVEESRPFPRRYDV 366
 QY 367 AIAEGHAYTLAAGQAQGAAPVAIYSTFLQSGYDQLIHDAVALQNLMDLFDALDRAGLVP 426
 DB 367 AIAEGHAYTLAAGQAQGAAPVAIYSTFLQSGYDQLIHDAVALQNLMDLFDALDRAGLVP 426
 QY 427 DGTTHAGADYSIMRCIPNMLTMAPADENECQMLTTGQHH-GEPASTRYPPKGPQAAI 485
 DB 427 DGTTHAGADYSIMRCIPNMLTMAPADENECQMLTTGQHH-GEPASTRYPPKGPQAAI 486
 QY 486 DPLTLAIEIGKAEVRHSGRIALANGSWYTPAVENGQOLGATVVMNMFVKPFQDVALIE 545
 DB 487 TP-LKCLPIKGIKVGKRGKRLALNFTGLMPPAKVAESLANTLVDMKFTVPLDALLIE 545
 QY 546 LAETHDVFTVEENVVIAAGAGSAINTFLOAKVLMFVNCIGLPDRFVBOGSREELSLVIG 605
 DB 546 MAASHALVYVENENALIMGAGSGVNEVILMAHRAKVPVNLIGLPDFFIPIQITQEEWRALIG 605
 QY 606 LBSGLIATIEQCA 620

Db 606 LBAAGWEAKIKAWILA 620

RESULT 12

US-10-369-493-15911
 : Sequence 15911, Application US/10369493
 : Publication No. US20030233675A1
 : GENERAL INFORMATION:
 : APPLICANT: Cao, Yongwei
 : APPLICANT: Hinkle, Gregory J.
 : APPLICANT: Slater, Steven C.
 : APPLICANT: Goldman, Barry S.
 : APPLICANT: Chen, Xianfeng
 : TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 : FILE REFERENCE: 38-10(52052)B
 : CURRENT APPLICATION NUMBER: US/10/369,493
 : PRIORITY FILING DATE: 2003-02-28
 : PRIOR APPLICATION NUMBER: US 60/360,039
 : NUMBER OF SEQ ID NOS: 47374
 : SEQ ID NO 15911
 : LENGTH: 623
 : TYPE: PRT
 : ORGANISM: Xanthomonas campestris
 : US-10-369-493-15911

Query Match 62.0%; Score 1992.5; DB 12; Length 623;
 Best Local Similarity 61.4%; Pred. No. 1.7e-181;
 Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

QY 4 TTDYPLLNKIHHPADIRALSKOQLQDLADVSGYLTHTVSGSGHFAAGLGVETVALHY 63
 DB 5 STYVPLRSIQTPDDIRTEBEADLTVADELPAVLTESYKSGHFAAGLGVETVALHY 64
 QY 64 YVNTPTDOLVMDVGHQAVPHKILTGKREMPFIRTLGVSAPFARDESSEYDAFGVGSSTSI 123
 DB 65 YVNTPTDOLVMDVGHQAVPHKILTGKREMPFIRTLGVSAPFARDESSEYDAFGVGSSTSI 124
 QY 124 TSISALGMAIALSOLGEGEDKKNVAILIGDGSITGGMAVEANAGDVMANLVLINDND 181
 DB 125 TSISALGMAIALSOLGEGEDKKNVAILIGDGSITGGMAVEANAGDVMANLVLINDND 184
 QY 182 MSISPVGAMNNYLTQVLSKFSYSSVRESKKALA--KMPSVWEILAKTEEHYKGVITVGTFFELG 239
 DB 185 MSISPVGAMNNYLTQVLSKFSYSSVRESKKALA--KMPSVWEILAKTEEHYKGVITVGTFFELG 244
 QY 240 TLFEEIGFNYGPIIDGHDVEMVSTLENLKDILTGPFVLTHTVTKKGVAFAEKDPLAVGVP 299
 DB 245 TLFEEIGFNYGPIIDGHDVEMVSTLENLKDILTGPFVLTHTVTKKGVAFAEKDPLAVGVP 304
 QY 300 VPAFDPTKDLPLKAPSPHPTTYEVGKRLCDMAADDERLGLITPAMREGSLVFSQTFPRRYDV 359
 DB 305 VGPFDPSKGLVAAAG-AKPTTYDVPSDVCMAAADPTDLITPAMREGSGMVEESRPFPRRYDV 363
 QY 360 PNRYFPAVALAEGHAYTLAAGQAQGAAPVAIYSTFLQSGYDQLIHDAVALQNLMDLFDALD 419
 DB 364 PNRYFPAVALAEGHAYTLAAGQAQGAAPVAIYSTFLQSGYDQLIHDAVALQNLMDLFDALD 423
 QY 420 RAGLVGDDGFTTHAGADYSIMRCIPNMLTMAPADENECQMLTTGQHH-GEPASTRYPPKGPQAAI 479
 DB 424 RAGLVGDDGFTTHAGADYSIMRCIPNMLTMAPADENECQMLTTGQHH-GEPASTRYPPKGPQAAI 483
 QY 480 GPGAAIDPLTLAIEIGKAEVRHSGRIALANGSWYTPAVENGQOLGATVVMNMFVKPFQDVALIE 539
 DB 484 GPGTALDASITLPIKGIKVGKRGKRLALNFTGLMPPAKVAESLANTLVDMKFTVPLDALLIE 543
 QY 540 QALVLEIARTHDVFTVEENVVIAAGAGSAINTFLOAKVLMFVNCIGLPDRFVBOGSREELSLVIG 599
 DB 544 KAMLEIARTHDVFTVEENVVIAAGAGSAINTFLOAKVLMFVNCIGLPDRFVBOGSREELSLVIG 603
 QY 600 LBSGLIATIEQCA 615

Db 604 LIAEAGIDGAGIRAAV 619

RESULT 13
US-10-369-493-16281
Sequence 16281, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 16281
LENGTH: 623
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-16281

Query Match 62.0%; Score 1992.5; DB 12; Length 623;
Best Local Similarity 61.4%; Pred. No. 1.7e-181;
Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

QY 4 TTDYPLKXHTPADIRALSKDQQLADEVGYTHTVSISGGHFAAGLTVETVALH 63
DB 5 STRYPLRSITQTPDLRTFEENDLTAVADELRAVLESVSGSGHFAAGLTVETVALH 64
QY 64 YVNTPTVDQVWDVGHQAVPHKILTGKREKMTITRTGGVSAFPARDESEYDAFGVSHS 123
DB 65 YVGTPTVDQVWDVGHQAVPHKILTGKREKMTITRTGGVSAFPARDESEYDAFGVSHS 124
QY 124 TSISALAGMALASQURGEDKXVAVIIGGSGITGMAVYANAHAG--DYNAALVYLINDND 181
DB 125 TSISALAGMALAORNGDRKVAVIGGAMTAGVYALAHAGGMDPEENLVYLINDNR 184
QY 182 MSISPPVGMANNYLTKVLSKFSVSRESESKKALA--KPSVWEIARKTEEHVKGMTVPG 239
DB 185 MSISEAVGSLTMDLRASGSRILNAREGCKKILCDKKNPTAFVRWEHWMGMFVPS 244
QY 240 TTFEBLGENYFGPIGDHVENLVSTLENLKDLTGVEFLHVVTKGKGTAPEKOPLAYHG 299
DB 245 TTFEBWGFYTPGPIGDHLPISLVGALKTKLTKLKGQILHVVITTKGKGTELAEGDQIGYHA 304
QY 300 VPAFPTKDFLKPAAKPSHPPTTEYFGWMLCDMAAODERLGTFPAMEBSGLVFEPSQKF 359
DB 305 VGPFPDSKGLVAKAG-AKGPYTTDVFSWVCMMAADPKMLVTPPAMEBSGLVRFSEY 363
QY 360 PNEYFDVAIAEQHAATVLAAGACOGAKPVVAIYSTFLOGVYDQILHVAALNLDMLFALD 419
DB 364 PORYFDVAIAEQHAATVLAAGACOGAKPVVAIYSTFLOGVYDQILHVAALNLDMLFALD 423
QY 420 RAGLVGPGPTTAGAFDVSYRCIPNMLIMAPADNECROMLTTGFOHNGASVRYPRGK 479
DB 424 RGVGVGPGDGTAGKGLDLSFLKCVPHMYVMAADABCRQMLTTGLRSGHAAIRYPRGT 483
QY 480 GPGALDPTLTALISGKAEVRHSGRIALANGSNVTPAVBAGKOLGATVNNKPFVCFPD 539
DB 484 GGTMLDASLTTLPIGKQOLRHSGARIALLGATVDAABAAGVREGILGTVNNMFVFXELD 543
QY 540 QALVLEIARTHDVFTVVEENVYAGAGSAINFTLOAQVYLMPCVNCIGLPDFVQGSREE 599
DB 544 KMLLELAKCHAPFISIEDNVVAGAGSVSELLNAESVIMPMHLGILPDFQHAASRED 603
QY 600 LSLVGLDSKGIILATI 615

Db 604 LIAEAGIDGAGIRAAV 619

RESULT 14
US-10-369-493-15540
Sequence 15540, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15540
LENGTH: 653
TYPE: PRT
ORGANISM: Xanthomonas campestris
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(653)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-15540

Query Match 62.0%; Score 1992.5; DB 12; Length 653;
Best Local Similarity 61.4%; Pred. No. 1.8e-181;
Matches 378; Conservative 91; Mismatches 142; Indels 5; Gaps 3;

QY 4 TTDYPLKXHTPADIRALSKDQQLADEVGYTHTVSISGGHFAAGLTVETVALH 63
DB 27 STRYPLRSITQTPDLRTFEENDLTAVADELRAVLESVSGSGHFAAGLTVETVALH 96
QY 64 YVNTPTVDQVWDVGHQAVPHKILTGKREKMTITRTGGVSAFPARDESEYDAFGVSHS 123
DB 87 YVGTPTVDQVWDVGHQAVPHKILTGKREKMTITRTGGVSAFPARDESEYDAFGVSHS 146
QY 124 TSISALAGMALASQURGEDKXVAVIIGGSGITGMAVYANAHAG--DYNAALVYLINDND 181
DB 147 TSISALAGMALAORNGDRKVAVIGGAMTAGVYALAHAGGMDPEENLVYLINDNR 206
QY 182 MSISPPVGMANNYLTKVLSKFSVSRESESKKALA--KPSVWEIARKTEEHVKGMTVPG 239
DB 207 MSISEAVGSLTMDLRASGSRILNAREGCKKILCDKKNPTAFVRWEHWMGMFVPS 266
QY 240 TTFEBLGENYFGPIGDHVENLVSTLENLKDLTGVEFLHVVTKGKGTAPEKOPLAYHG 299
DB 267 TTFEBWGFYTPGPIGDHLPISLVGALKTKLTKLKGQILHVVITTKGKGTELAEGDQIGYHA 326
QY 300 VPAFPTKDFLKPAAKPSHPPTTEYFGWMLCDMAAODERLGTFPAMEBSGLVFEPSQKF 359
DB 327 VGPFPDSKGLVAKAG-AKGPYTTDVFSWVCMMAADPKMLVTPPAMEBSGLVRFSEY 385
QY 360 PNEYFDVAIAEQHAATVLAAGACOGAKPVVAIYSTFLOGVYDQILHVAALNLDMLFALD 419
DB 386 PORYFDVAIAEQHAATVLAAGACOGAKPVVAIYSTFLOGVYDQILHVAALNLDMLFALD 445
QY 420 RAGLVGPGPTTAGAFDVSYRCIPNMLIMAPADNECROMLTTGFOHNGASVRYPRGK 479
DB 446 RGVGVGPGDGTAGKGLDLSFLKCVPHMYVMAADABCRQMLTTGLRSGHAAIRYPRGT 505
QY 480 GPGALDPTLTALISGKAEVRHSGRIALANGSNVTPAVBAGKOLGATVNNKPFVCFPD 539
DB 506 GGTMLDASLTTLPIGKQOLRHSGARIALLGATVDAABAAGVREGILGTVNNMFVFXELD 565
QY 540 QALVLEIARTHDVFTVVEENVYAGAGSAINFTLOAQVYLMPCVNCIGLPDFVQGSREE 599

Db 566 KAMLELAKCHEAFISIDNVAGAGSGVELLNAESVLPMLHGLPDSFQIHAERD 625
Qy 600 LLSLVGDSKGLIATI 615
Db 626 LIAEXGIDQAGIRANV 641

RESULT 15

US-10-369-493-7428
; Sequence 7428, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7428
; LENGTH: 619
; TYPE: PR1
; ORGANISM: Burkholderia cepacia
US-10-369-493-7428

Query Match 61.9%; Score 1990.5; DB 12; Length 619;

Best Local Similarity 61.7%; Pred. No. 2.6e-181;

Matches 381; Conservative 89; Mismatches 141; Indels 7; Gaps 3;

Qy 4 TTDVPLKNIHTPADIRALSKDQLOQLADEVRGYLTHPTVSGHFAAGLTVELVALH 63
Db 3 TTMVPLKNIHTPADIRALSKDQLOQLADEVRGYLTHPTVSGHFAAGLTVELVALH 62
Qy 64 YVFNTPVDQVMDVGHQAYPHKILTKREKMPITRTLGVSAPFAPDSEYDAFGVHSS 123
Db 63 YVFNTPVDQVMDVGHQAYPHKILTKREKMPITRTLGVSAPFAPDSEYDAFGVHSS 122
Qy 124 TSISAALGMAISQLRGDKKXVATIGSGSITGGAYEAMNHAG-DVNAALVTINDNM 182
Db 123 TSISAALGMAISQLRGDKKXVATIGSGSITGGAYEAMNHAG-DVNAALVTINDNM 182
Qy 183 SISPPVGMNNTLYTVLSSKPYSSVREBSKALAMPSEWELARKTEBHVKGMIVPGLF 242
Db 182 SISPPVGMNNTLYTVLSSKPYSSVREBSKALAMPSEWELARKTEBHVKGMIVPGLF 242
Qy 243 BELGNNYSGPIRGHVMVSTLENKDLGVPVLYVTKKGYAPAEKDLAYHGVPA 302
Db 242 BELGNNYSGPIRGHVMVSTLENKDLGVPVLYVTKKGYAPAEKDLAYHGVPA 302
Qy 243 BELGNNYSGPIRGHVMVSTLENKDLGVPVLYVTKKGYAPAEKDLAYHGVPA 302
Db 242 BELGNNYSGPIRGHVMVSTLENKDLGVPVLYVTKKGYAPAEKDLAYHGVPA 302
Qy 303 PDPTDPLPKAAPSPHPTVTEVFGRMLCDMAODERLIGITPAMBSGGLVFSQKPPNR 362
Db 302 PDPTDPLPKAAPSPHPTVTEVFGRMLCDMAODERLIGITPAMBSGGLVFSQKPPNR 362
Qy 303 FMPAAGIKPAATPS-KMYTYGVFGEMLDAAELDARVIGITPAMBSGGLVFSQKPPNR 361
Db 302 FMPAAGIKPAATPS-KMYTYGVFGEMLDAAELDARVIGITPAMBSGGLVFSQKPPNR 361
Qy 363 YFDVVAIAEGHATVLAAGQACGAKPVAYISTFLORGVDLIHVAALQNLDMFLALDRAG 422
Db 362 YFDVVAIAEGHATVLAAGQACGAKPVAYISTFLORGVDLIHVAALQNLDMFLALDRAG 422
Qy 423 LVGPGPTAGAFDYSYMRCTPMLIMAPADENECRQMLTTFQHGHPASVYPRGKPG 482
Db 422 LVGPGPTAGAFDYSYMRCTPMLIMAPADENECRQMLTTFQHGHPASVYPRGKPG 482
Qy 483 AALDPTLALTEGKAEVRRH-----GSRIALIAGSMVTPAVERAGKQIGATVVMRFYKP 537
Db 482 AALDPTLALTEGKAEVRRH-----GSRIALIAGSMVTPAVERAGKQIGATVVMRFYKP 537
Qy 538 POGALVILBARTHDVFTVVEENVVAGGASALNTFLQAKVLMPCNIGLPDRFVEGSR 597
Db 537 POGALVILBARTHDVFTVVEENVVAGGASALNTFLQAKVLMPCNIGLPDRFVEGSR 597

Db 542 LDADLVRLAETHDAIVTEBGCWGGAGSACVEALLASGVTRPVQLGLPDRFIDHGP 601
Qy 598 BELLSLVGDSKGLIATI 615
Db 602 AKLLAAGCIDAVGITKSI 619

Search completed: January 29, 2004, 16:21:14
Job time : 44.325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:37:58 ; Search time 39 Seconds
(without alignments)

2328.695 Million cell updates/sec

Title: US-09-941-947a-2

Perfect score: 2284

Sequence: 1 DVVTWPHLTADIRFCWTF.....FNMELSRIGQPKGKRYEVSH 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2284	100.0	437	10	US-09-934-901-16
2	2284	100.0	437	10	US-09-934-868-6
3	2284	100.0	437	11	US-09-941-947a-2
4	2284	100.0	437	12	US-10-320-874-16
5	2284	100.0	437	16	US-10-320-924-16
6	659	28.9	184	12	US-10-369-493-14348
7	306.5	13.4	350	12	US-10-369-493-8108
8	294	12.9	342	15	US-10-156-761-13610
9	284.5	12.5	341	15	US-10-156-761-14652
10	269.5	11.8	341	15	US-10-156-761-10359
11	261	11.4	345	12	US-10-369-493-19261
12	253	11.1	319	12	US-10-369-493-2905
13	250.5	11.0	357	12	US-10-369-493-9669
14	246.5	10.8	987	12	US-10-369-493-21998
15	240.5	10.5	320	9	US-09-815-242-13785

16	239.5	10.5	318	12	US-10-369-493-9652	Sequence 9652, Ap
17	238.5	10.4	320	9	US-09-815-242-10422	Sequence 10422, A
18	238.5	10.4	320	12	US-10-369-493-23638	Sequence 23638, A
19	229.5	10.0	295	12	US-10-369-493-8855	Sequence 8855, Ap
20	227.5	10.0	449	12	US-10-369-493-8876	Sequence 8876, Ap
21	226.5	9.9	320	9	US-09-815-242-11721	Sequence 11721, A
22	224.5	9.8	334	12	US-10-369-493-515	Sequence 515, Ap
23	223.5	9.8	319	12	US-10-369-493-17399	Sequence 17399, A
24	222.5	9.7	322	9	US-09-815-242-5425	Sequence 5425, Ap
25	222.5	9.7	322	9	US-09-815-242-12619	Sequence 12619, A
26	222.5	9.7	322	9	US-09-815-242-12907	Sequence 12907, A
27	222.5	9.7	360	12	US-10-369-493-19198	Sequence 19198, A
28	222	9.7	322	12	US-10-369-493-10132	Sequence 10132, A
29	222	9.7	340	12	US-10-369-493-18431	Sequence 18431, A
30	221	9.7	335	9	US-09-815-242-13346	Sequence 13346, A
31	218.5	9.6	320	9	US-09-815-242-10649	Sequence 10649, A
32	215	9.4	321	12	US-10-369-493-21210	Sequence 21210, A
33	213.5	9.3	321	12	US-10-369-493-128	Sequence 128, App
34	213	9.3	321	12	US-10-369-493-479	Sequence 479, App
35	207	9.1	959	12	US-10-369-493-1903	Sequence 1903, Ap
36	202.5	8.9	321	9	US-09-815-242-11162	Sequence 11162, A
37	201.5	8.8	361	12	US-10-369-493-2625	Sequence 2625, Ap
38	201	8.8	343	9	US-09-887-054-2	Sequence 2, Appl
39	201	8.8	343	15	US-10-098-626-2	Sequence 2, Appl
40	201	8.8	357	10	US-09-738-626-4883	Sequence 4883, Ap
41	200	8.8	776	15	US-10-305-342-5	Sequence 5, Appl
42	198	8.7	419	12	US-10-369-493-2919	Sequence 2919, Ap
43	196	8.6	321	12	US-10-369-493-10434	Sequence 10434, A
44	196	8.6	828	12	US-10-369-493-6730	Sequence 6730, Ap
45	193	8.5	942	12	US-10-369-493-2497	Sequence 2497, Ap

ALIGNMENTS

RESULT 1
US-09-934-901-16
Sequence 16, Application US/09934901
Patient No. US20020110885A1
GENERAL INFORMATION
APPLICANT: scifac, Matthews
APPLICANT: Odum, J. Martin
APPLICANT: No. US20020110885Altom, Kelley C.
APPLICANT: Ye, Rick
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: C14619 US NA
CURRENT APPLICATION NUMBER: US/09/934, 901
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229, 906
PRIOR FILING DATE: September 1, 2000
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 437
TYPE: PRT
ORGANISM: METHYLOMONAS SP.
US-09-934-901-16

Query Match 100.0%; Score 2284; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.4e-223;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DVVTWPHLTADIRFCWTF...FNMELSRIGQPKGKRYEVSH 60
DB 1 DVVTWPHLTADIRFCWTF...FNMELSRIGQPKGKRYEVSH 60
QY 61 IDPFEIICRGYKGLLDSSYPVTAVERKKGAVLQRFQSGVIGNSRVLTAVKQCVKR 120
DB 61 IDPFEIICRGYKGLLDSSYPVTAVERKKGAVLQRFQSGVIGNSRVLTAVKQCVKR 120
QY 121 GLVTEGEDPQKVAADQVLKQGVDIHTIGDDTNTAAADLAATLARNRYGLTYIGPKTV 180
DB 121 GLVTEGEDPQKVAADQVLKQGVDIHTIGDDTNTAAADLAATLARNRYGLTYIGPKTV 180

QY 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
DB 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
QY 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
DB 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
QY 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
DB 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
QY 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
DB 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
QY 421 MLESEIGOPKGGKXEVSH 437
DB 421 MLESEIGOPKGGKXEVSH 437

RESULT 2

US-09-934-868-6
Sequence 6, Application US/09934868
Patent No. US20020137190A1
GENERAL INFORMATION:
APPLICANT: Koffas, Mattheos
APPLICANT: Odum, James M.
APPLICANT: Schenzle, Andreas J.
TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: C1156 US NA
CURRENT APPLICATION NUMBER: US/09/934,868
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 437
TYPE: PRT
ORGANISM: METHYLIOMONAS SP.
US-09-934-868-6

Query Match 100.0%; Score 2284; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.4e-223;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWMPYHLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60
DB 1 DVTWMPYHLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60
QY 61 IDPSIIEIICRYGGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120
DB 61 IDPSIIEIICRYGGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120
QY 121 GLVEGEDPOKVAADQVLKQGVDI LHTTIGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180
DB 121 GLVEGEDPOKVAADQVLKQGVDI LHTTIGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180
QY 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
DB 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
QY 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
DB 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
QY 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
DB 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
QY 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
DB 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420

DB 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
QY 421 MLESEIGOPKGGKXEVSH 437
DB 421 MLESEIGOPKGGKXEVSH 437

RESULT 3

US-09-941-947a-2
Sequence 2, Application US/09941947a
Publication No. US20030003528A1
GENERAL INFORMATION:
APPLICANT: Bzostowicz, Patricia C.
APPLICANT: Cheng, Qiong
APPLICANT: DiCosimo, Deana J.
APPLICANT: Koffas, Mattheos
APPLICANT: Miller, Edward S. Jr.
APPLICANT: Odum, J. Martin
APPLICANT: Picataggio, Steve
APPLICANT: Rouviere, Pierre E.
TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
FILE REFERENCE: C11903 US NA
CURRENT APPLICATION NUMBER: US/09/941,947a
CURRENT FILING DATE: 2001-09-01
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,858
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 437
TYPE: PRT
ORGANISM: Methyliomonas 16a
US-09-941-947a-2

Query Match 100.0%; Score 2284; DB 11; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.4e-223;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWMPYHLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60
DB 1 DVTWMPYHLTADIRFCMFPLNFYTLNKPCKKVALITAGGLAPCLNSAIGSLIRYTE 60
QY 61 IDPSIIEIICRYGGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120
DB 61 IDPSIIEIICRYGGYKGLLDSDYPVTAEVKKAQVLOREFGSVI GNSRVLTNVKDCVKR 120
QY 121 GLVEGEDPOKVAADQVLKQGVDI LHTTIGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180
DB 121 GLVEGEDPOKVAADQVLKQGVDI LHTTIGGDDNTAAADLAFLAENNYGLTVIGLPKTV 180
QY 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
DB 181 DNDVFPKOSIGANTAAOGARFENNVAAENNANPRMLIVHEVWGNCGMLTAATAOEYR 240
QY 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
DB 241 KLDRAEWLPELGLTRESYEVHAFVPEMAIDLEAEKRLREVMKRYDCVNTFVSEGAGV 300
QY 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
DB 301 EAIYAEMOAKGOEYPRDAFGHILKIDAVNPGKMGEOPOAOMIGAEKTLVOKSGYFAPASAS 360
QY 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
DB 361 NVDMRLIKSCADLAIVACAFRRBSGVI GHEDNGNVLRAIEFPRIKGGKPFNIDTWMFNS 420
QY 421 MLESEIGOPKGGKXEVSH 437
DB 421 MLESEIGOPKGGKXEVSH 437

RESULT 4
 US-10-320-874-16
 ; Sequence 16, Application US/10320874
 ; Publication No. US20030138909A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Koffas, Metcheos
 ; APPLICANT: Koffas, Metcheos
 ; APPLICANT: No. US20030138909A1, Kelley C.
 ; APPLICANT: Ye, Rick
 ; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
 ; FILE REFERENCE: CL1619 US NA
 ; CURRENT APPLICATION NUMBER: US/10/320,874
 ; CURRENT FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: US/09/934,901
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/229,906
 ; PRIOR FILING DATE: September 1, 2000
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 16
 ; LENGTH: 437
 ; TYPE: PRT
 ; ORGANISM: METHYLOMONAS SP.
 US-10-320-874-16

Query Match 100.0%; Score 2284; DB 12; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1,4e-223;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DVTWPHLTDADIRCHFWFLNFPYTLNPKKVAIITAGLAPCLNSAIGSLIERYTE	60
DB	1	DVTWPHLTDADIRCHFWFLNFPYTLNPKKVAIITAGLAPCLNSAIGSLIERYTE	60
QY	61	IDPSIEIICRGYKGLLDGSDPYTAVERKAGVLRFGSVIGNSRVKTIVKDCVCR	120
DB	61	IDPSIEIICRGYKGLLDGSDPYTAVERKAGVLRFGSVIGNSRVKTIVKDCVCR	120
QY	121	GLVKEGDPORVADOLVKDGVLDLHTIGGDDTAAADLAFLARNNGITVIGLRTV	180
DB	121	GLVKEGDPORVADOLVKDGVLDLHTIGGDDTAAADLAFLARNNGITVIGLRTV	180
QY	181	DNDVFPRIKOSIGAMTAAGARFYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOBYR	240
DB	181	DNDVFPRIKOSIGAMTAAGARFYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOBYR	240
QY	241	KLDRAEWLPELGLTRSEYVHAFFVPEMAIDLEAKRLREVNDKVDVNIIVSEAGV	300
DB	241	KLDRAEWLPELGLTRSEYVHAFFVPEMAIDLEAKRLREVNDKVDVNIIVSEAGV	300
QY	301	EAIYAEKQAEVPRDAFGHIKLDVAVNGKMFGEOPQMTGAETLVOKSGYFARASAS	360
DB	301	EAIYAEKQAEVPRDAFGHIKLDVAVNGKMFGEOPQMTGAETLVOKSGYFARASAS	360
QY	361	NVDWRLIKSCADLAABCAFRRESGVI GHEDDNGVLRALIEPRIKGKPNIDTWMNS	420
DB	361	NVDWRLIKSCADLAABCAFRRESGVI GHEDDNGVLRALIEPRIKGKPNIDTWMNS	420
QY	421	MLSEIGOPKRGKVEVSH 437	
DB	421	MLSEIGOPKRGKVEVSH 437	

RESULT 5
 US-10-320-924-16
 ; Sequence 16, Application US/10320924
 ; Publication No. US20030129721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Koffas, Metcheos
 ; APPLICANT: Koffas, Metcheos
 ; APPLICANT: No. US20030129721A1, Kelley C.
 ; APPLICANT: Ye, Rick
 ; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA
 ; CURRENT APPLICATION NUMBER: US/10/320,924
 ; CURRENT FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: US/09/934,901
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/229,906
 ; PRIOR FILING DATE: September 1, 2000
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 16
 ; LENGTH: 437
 ; TYPE: PRT
 ; ORGANISM: METHYLOMONAS SP.
 US-10-320-924-16

Query Match 100.0%; Score 2284; DB 16; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1,4e-223;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DVTWPHLTDADIRCHFWFLNFPYTLNPKKVAIITAGLAPCLNSAIGSLIERYTE	60
DB	1	DVTWPHLTDADIRCHFWFLNFPYTLNPKKVAIITAGLAPCLNSAIGSLIERYTE	60
QY	61	IDPSIEIICRGYKGLLDGSDPYTAVERKAGVLRFGSVIGNSRVKTIVKDCVCR	120
DB	61	IDPSIEIICRGYKGLLDGSDPYTAVERKAGVLRFGSVIGNSRVKTIVKDCVCR	120
QY	121	GLVKEGDPORVADOLVKDGVLDLHTIGGDDTAAADLAFLARNNGITVIGLRTV	180
DB	121	GLVKEGDPORVADOLVKDGVLDLHTIGGDDTAAADLAFLARNNGITVIGLRTV	180
QY	181	DNDVFPRIKOSIGAMTAAGARFYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOBYR	240
DB	181	DNDVFPRIKOSIGAMTAAGARFYFMNVVAENNANPRLIYHEVWGRNCGMLTAATAOBYR	240
QY	241	KLDRAEWLPELGLTRSEYVHAFFVPEMAIDLEAKRLREVNDKVDVNIIVSEAGV	300
DB	241	KLDRAEWLPELGLTRSEYVHAFFVPEMAIDLEAKRLREVNDKVDVNIIVSEAGV	300
QY	301	EAIYAEKQAEVPRDAFGHIKLDVAVNGKMFGEOPQMTGAETLVOKSGYFARASAS	360
DB	301	EAIYAEKQAEVPRDAFGHIKLDVAVNGKMFGEOPQMTGAETLVOKSGYFARASAS	360
QY	361	NVDWRLIKSCADLAABCAFRRESGVI GHEDDNGVLRALIEPRIKGKPNIDTWMNS	420
DB	361	NVDWRLIKSCADLAABCAFRRESGVI GHEDDNGVLRALIEPRIKGKPNIDTWMNS	420
QY	421	MLSEIGOPKRGKVEVSH 437	
DB	421	MLSEIGOPKRGKVEVSH 437	

RESULT 6
 US-10-369-493-14348
 ; Sequence 14348, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 14348
 ; LENGTH: 184
 ; TYPE: PRT

ORGANISM: Agrobacterium tumefaciens
US-10-369-493-13348

Query Match 28.9%; Score 659; DB 12; Length 184;
Best Local Similarity 67.9%; Pred. No. 1.1e-58;
Matches 125; Conservative 27; Mismatches 32; Indels 0; Gaps 0;

QY 33 KXVAILTAGLAPCLNSAIGSLERTEIDPSIEII CYGKGKLLGDSVPVTAVERK 92
DB 1 QXVAMLTGGLAPCLNSAVGSLERISDLPEZIDITAYSGVGLGERITTKOMREK 60
QY 93 AGVLOREFGSVIGNSRYKLTNVKDCVKEGLVKEGEDPQKVAADQVGLHTIGDD 152
DB 61 AHIHRYGSGPIGSRVKLTNAADCARGLVKEGDNPRLVAERLAAOCITILHTIGDD 120
QY 153 TMTAAADLAFLAANNYGLTVIGLPTVNDVFPRIKOSLGAMTAAGARVPMVAENK 212
DB 121 TMTAAADLAFLAANGVGLTVIGLPTVNDVFPRIKOSLGAMTAAGARVPMVAENK 180
QY 213 ANPR 216
DB 181 ANPR 184

RESULT 7

US-10-369-493-8108
Sequence 8108, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 8108
LENGTH: 350
TYPE: PRT
ORGANISM: Thermobifida fusca
US-10-369-493-8108

Query Match 13.4%; Score 306.5; DB 12; Length 350;
Best Local Similarity 26.5%; Pred. No. 2.2e-22;
Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNYFTLANKPKKVALITAGLAPCLNSAIGSLERTEIDPSIEII CYGKGKLLGDS 82
DB 4 FKEATQK---RQVGLTGSGDPCGLNAVIRAVVRKGLK-EYGEFVGFDGMRGPLEGUT 58
QY 83 YPVTAVERKXKAGVQRFQSGSVIGNSRYKLTNVKDCVKEGLVKEGEDPQKVAADQVGL 142
DB 59 MFLIEAVR-GILPR-GGITLSSRTNLKIBGVR-----VDNMAALAV 103
QY 143 DILHTIGGDDTNTAAADLAFLAANNYGLTVIGLPTVNDVFPRIKOSLGAMTAAGAR 202
DB 104 DALVAIGGEDTLGVAROL-----HDGVAVVGVPKTIDNDLANATYTFGFDVAVNIATE 157
QY 203 YFNVVAVENNANPKYLLVHEVMGRNCGMLT-----ATAQRYKLLDRAEMLPELGITRE 257
DB 158 AIDRLHTTASHSRHALVY-EVGGHAGKIALAHAGMAGA----- 195
QY 258 SEYVAVFVPEMAIDLEABARLEAVMDKVCNIPVSEAGVEALVAENQAKGVEPRD 317
DB 196 ---NVILIFRPRPDIDVAVYIESRFTNVAPIIVABAHK---EGQLTAGAERD 247
QY 318 AFGHIKLDVAVNPGMFGGQFQMGIGAEITLVQKS---GYFAASASNVDMRLIKSCADL 374

DB 248 SFGVRLVIGI-----GQRLAEIEARTKEARSVVLGHVQRGTPSAFDRVLAIRLGLH 301
QY 375 AVECAFRRESGVIGHDEDNQNV--LRAISFPR 405
DB 302 ALTRV-----HDKDFKQVALLGTRIVRV 325

RESULT 8

US-10-156-761-13610
Sequence 13610, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OMIYA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13610
LENGTH: 342
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-13610

Query Match 12.9%; Score 294; DB 15; Length 342;
Best Local Similarity 26.5%; Pred. No. 4e-21;
Matches 101; Conservative 66; Mismatches 140; Indels 74; Gaps 17;

QY 34 KVALITGGLAPCLNSAIGSLERTEIDPSIEII CYGKGKLLGDSVPVTAVERK 93
DB 2 RVGLTGSGDPCGLNAVIRKGVQ-EYGYDVGFDGMRGPLEGDA--VALDIPAR 58
QY 94 GVLQRFQSGSVIGNSRYKLTNVKDCVKEGLVKEGEDPQKVAADQVGLHTIGDDT 153
DB 59 GILPR-GGITLSSRTNLKIBGVR-----NLAKQVDALIAIGEDT 105
QY 154 MTAADLAFLAANNYGLTVIGLPTVNDVFPRIKOSLGAMTAAGARV--MNVVAEN 211
DB 106 LGVAARLT-----DEYGVAVVGVPKTIDNDLSATDYTFGFDVAVNIATEIDRLHTTALS 160
QY 212 NAMPKLTIVHVMGRNCGMLTAATAQRYKLLDRAEMLPELGITRESYEVHAVFVPMAL 271
DB 161 HRR--VLVGVVGRHAGWLAHIS-----GLAGCA--NVILIEQRF 197
QY 272 DLEAARLEAVMDKVCNIPVSEAGVEALVAENQAKGVEPR---DAFHTLDAY 327
DB 198 DVDQVCAVYISRFQSTAYAPIVVAEGA-----MFKQGMWLKQESLDSFQHVLSGV 249
QY 328 NRGKMPGQFQMGIGAE-KTLVQKSGYFAASASNVDMRLIKSCADLAVECAFRRESGV 386
DB 250 --GEMLAKEIKERKGEAKRTTV--LGHVQGTGPSAFDRVLAIRFGHIAEV----- 258
QY 387 IGHDEDNQNV--LRAISFPR 405
DB 299 --RDGDFKQVALLGTRIVRV 317

RESULT 9

US-10-156-761-14652
Sequence 14652, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

```

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, KASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14652
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-14652

```

```

Query Match      12.5% Score 284.5; DB 15; Length 341;
Best Local Similarity 27.5%; Pred. No. 3.7e-20;
Matches 97; Conservative 62; Mismatches 143; Indels 51; Gaps 11;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEITCYRGYKGLLDSDYPTAEVAKKA 93
DB 2 RIGVLTGGDCPGLAIVRSVYVHR-AVVDHGDDEVIGFRDQKGLLECDY--LKDLDLAVS 58
QY 94 GVLORFGGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQLVKQGVILHTIGSDDT 153
DB 59 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAAREFDLALIPIGSGGT 105
QY 154 NTAADLAAFLARNNGITVIGLEPTVNDVFPKOSIGAMTAABQCARFMMVVAENNA 213
DB 106 LKAARLT-----SDAGLPVGVPEKTDINDIAVDVTFGPTAVGAVATEADRLKTTAES 159
QY 214 NPMRLIVHVMGRCNGMTLTAATQERYKLDRAEWLPELGLTRSYEVAHAFVPEMAIDL 273
DB 160 HORVWVW-EVMGRHAGWIAL-----ESGMAAGA--HGCTLERPPDP 198
QY 274 EAAARLREVMDC-----VDCVNI FVSGAGVEALVAEMQKGOEVPDPAFGITKLDVN 328
DB 199 ARLVWQVEERFARGKKPAVLC-----VAEGA--HPVQSGMDYSHGAI--DQGHRRFOGI 249
QY 329 PKWFGEPQOMIGAE--TLVQKSGYFARASAVNDMLIKSCADLAVECAFRREG 385
DB 250 -CTALAAELERLIGKEARPV--LGHVORGSTPTAYDRLATRGMTAVAEAAHRRDPT 304
QY 333 FGEOPQOMIGAEKTLVQKSGYFARASAVNDMLIKSCADLAVECAFRREG 385
DB 253 LSELLERLIGKEARPV--LGHVORGSTPTAYDRLATRGMTAVAEAAHRRDPT 304

```

RESULT 10

```

; Sequence 10359, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, KASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10359

```

```

; LENGTH: 341
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-10359

```

```

Query Match      11.8% Score 269.5; DB 15; Length 341;
Best Local Similarity 26.8%; Pred. No. 1.2e-18;
Matches 96; Conservative 59; Mismatches 142; Indels 61; Gaps 15;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEITCYRGYKGLLDSDYPTAEVAKKA 93
DB 2 RIGVLTGGDCPGLAIVRSVYVHR-VTVYGD-EVIGFEDQYAGLL--DGRYALDLAVS 58
QY 94 GVLORFGGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQLVKQGVILHTIGSDDT 153
DB 59 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAAREFDLALIPIGSGGT 105
QY 154 NTAADLAAFLARNNGITVIGLEPTVNDVFPKOSIGAMTAABQCARFMMVVAENNA 213
DB 106 LTAARLT-----SDAGLPVGVPEKTDINDIAVDVTFGPTAVGAVATEADRLKTTAES 159
QY 214 NPMRLIVHVMGRCNGMTLTAATQERYKLDRAEWLPELGLTRSYEVAHAFVPEMAIDL 273
DB 160 HORVWVW-EVMGRHAGWIAL-----ESGMAAGA--HGCTLERPPDP 198
QY 274 EAAARLREVMDC-----VDCVNI FVSGAGVEALVAEMQKGOEVPDPAFGITKLDVN 328
DB 199 ARLVWQVEERFARGKKPAVLC-----VAEGA--HPVQSGMDYSHGAI--DQGHRRFOGI 249
QY 329 PKWFGEPQOMIGAE--TLVQKSGYFARASAVNDMLIKSCADLAVECAFRREG 385
DB 250 -CTALAAELERLIGKEARPV--LGHVORGSTPTAYDRLATRGMTAVAEAAHRRDPT 304

```

RESULT 11

```

; Sequence 19261, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19261
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-10-369-493-19261

```

```

Query Match      11.4% Score 261; DB 12; Length 345;
Best Local Similarity 26.0%; Pred. No. 9.3e-18;
Matches 96; Conservative 66; Mismatches 141; Indels 66; Gaps 16;

```

```

QY 34 KVAITTAGLAPCLNSAIGSLIERYTEIDPSIEITCYRGYKGLLDSDYPTAEVAKKA 93
DB 2 KVAITLGGDCPGLAIVRSVYVHR--HGFEMGLRDMKLLLENNHRLRRET--TS 57
QY 94 GVLORFGGSVIGNSRV--KLTNVKDCVKGGLVEGEDPQKVAADQLVKQGVILHTIGG 150
DB 58 GILAR-GGTLIGSSSLQDRLREACEN-----AQDMAAREFDLALIPIGSGGT 105
QY 151 DDTAAADLAAFLARNNGITVIGLEPTVNDVFPKOSIGAMTAABQCARF--MNVV 208
DB 102 EGTLSAATM-----SQBLRIVGVPKTIDINDIATDPTFGDVAVALATAIDRLHST 155

```

```

QY 209 AENNANPRLIVHEVMGNCGLTAATAOERYKLDRAEWLPELGLTRESYEVHAFVPE 268
DB 156 AESH---KRVIVCVMGRHGMIT-----YAGIAGAD-----VILVEE 192
QY 269 MAIDLEAKRL--REVMKVDVNIIVSFGAGVEALVAMQAKGQVPRDAPGHTKIDA 326
DB 193 IPADLAKVAEHIQRRHAGRTFSI-VVAAGTRIKLSADQOGLVTSALDEKGRDRLG 251
QY 327 VNRKWFGECPAQMGIGAKTLVQKSGYFAPASASNDMDL-----IKSCADLAVECER 381
DB 252 V--GTILAHIERETGE-TRVSVLGHIOGAPVADRLVATRYGVHAC-DVVARGEERG 307
QY 382 RESGVIGHD 390
DB 308 KMAALRGND 316

```

```

RESULT 12
US-10-369-493-2905
; Sequence 2905, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2905
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2905

```

```

Query Match 11.1%; Score 253; DB 12; Length 319;
Best Local Similarity 26.6%; Pred. No. 5.4e-17;
Matches 95; Conservative 60; Mismatches 126; Indels 76; Gaps 14;
QY 33 KQVAILTAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKYGLLGDSPVTAVERKK 92
DB 2 KKLAVLTSGGDAFGNNAVAAY-RY-GVRQGLEVIQVRRGYSGLIDDF--VKLEVDV 57
QY 93 AGULORRGSVIGNSRYKLTNVVDQYKSGLVKEGEDPOKVAADQLVKQGVDLHTIGDD 152
DB 58 AGITER-GGITLTSCSEEFRT-----EEGRLEAKQIKQHIEGLVVIQEG 104
QY 153 TNPAADEAFLARNYGLTVIGLPTVDNDVFPFKOSLGAWTAAEQGARFTMNVAAEN 212
DB 105 SLF-----GAHLIYSEHKIPVVGIPATINDIGLTMCGIVDTCLNTVWDANQKIDTAS 159
QY 213 ANRPMILVHEVMGRNCGLTAATAOERYKLDRAEWLPELGLTRESYEVHAFVPEAID 272
DB 160 SHRAPIV-EVMRSHSGYIALMAG-----LVGTAB-----ATVPEIPVD 198
QY 273 LEAEAKRELV--MDKVDVNIIVSFGAGVEALVAMQAKGQVPRDAPGHTKIDAVNPG 330
DB 199 YSLADRLILEERRRGINSI-ILVAGASATVVA-----HLE----- 236
QY 331 KMGGBQPAQMGIGAKTLVQKSGYFAPASASNDMDLISKADLAVECERRESGVT 387
DB 237 -----YRIGYE-TRITLGHVORGSFPAFDRRLALSMGVEAVDALDDEVDM 284

```

RESULT 13
US-10-369-493-9669

```

; Sequence 9669, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9669
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Desulfotobacterium hafnense
US-10-369-493-9669

```

```

Query Match 11.0%; Score 250.5; DB 12; Length 357;
Best Local Similarity 27.4%; Pred. No. 1.1e-16;
Matches 101; Conservative 63; Mismatches 136; Indels 69; Gaps 17;
QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKYGLLGDSPVTAVERKK 93
DB 1 KKLITGGGCPGGLNIVRAAYK--TACRGIVAVIOGFRFRAVAGDREPL--ELKQVS 56
QY 94 GVLORFGSVIG-----NSRVKLTNVKQCVKQGLVKEGEDPOKVAADQLVKQGVDLHTI 148
DB 57 GLIPR-GGITLTGTNRDNPAPVPTOV-----GEERQVODRSALVDRKAREGIDALILAI 109
QY 149 GGDQNTAADLAFLARNYGLTVIGLPTVDNDVFPFKOSLGAWT--AAEQGARFTMNV 206
DB 110 GGGGSLIAWEFA-----KQGLKAVGVKRTINDLACTDITFGPQTAATATQDADLRH 163
QY 207 VVAENNANPRLIVHEVMGRNCGLTAATAOERYKLDRAEWLPELGLTRESYEVHAFV 266
DB 164 TTAESH---HRIMILVEMGRYAGWIAL-----YAGVAGAGD-----VILI 200
QY 267 PENALIDE--AEAKRREVMKVDVNIIVSFGA---GVEALVAMQAKGQVPRDAPGH 321
DB 201 PELPYQESIAEAVQRRARLGRFST-ILVAGAKPLGGDMV-ERTMSGRTDP----- 252
QY 322 IKLDVAVPGKWFGECPA---QMIGAKTLVQKSGYFAPASASNDMDLISKADLAVEC 378
DB 253 IKLGGI-----GAKLADLEKVTDMETVTVLGHLORGSPIDVDRVLTSTRYGVAVEA 306
QY 379 AFRRESGVT 387
DB 307 ALAGDPGMM 315

```

```

RESULT 14
US-10-369-493-21998
; Sequence 21998, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374

```

SEQ ID NO 21998
 LENGTH: 987
 TYPE: PR
 ORGANISM: Saccharomyces cerevisiae
 US-10-369-493-21998

Query Match 10.8%; Score 246.5; DB 12; Length 987;
 Best Local Similarity 23.9%; Pred. No. 1.4e-15;
 Matches 103; Conservative 68; Mismatches 135; Indels 125; Gaps 19;

QY 31 KKKKVALITAGGALPCNSAIGSLIERYTEIDPSIEIICRGYKGLLGSYVTAIVR 90
 DB 204 KKKKVALITAGGALPCNSAIGSLIERYTEIDPSIEIICRGYKGLLGSYVTAIVR 90
 QY 91 KKA-----GVICRFGSVIGNSRVKLTNVKDCVKGGLVKEGDDPQKVAADQLVKDVIL 145
 DB 257 KNAMEVDFRMGISE--GGTILGTA-----RSMFRRKRGSRQAGNLISQIDAL 303
 QY 146 HTIGDDTNTAADI-----AFLARNY-----GLTVIGLPTVNDVPEPI 187
 DB 304 VVCGGDSGLT--GADLFRHEWPSLVDELVAEGRFTEEVAPYKMLSTIVGLVGSIDNMSGT 362
 QY 188 KSLGAMTAEGGAFYFVAVVANNANPRMLIVHEVGRNCGMLT-----ATAQERYKL 243
 DB 363 DSTIGAVSALERICEMVDYIDATKSHSRFV--EVMGRHCGMLALMAGIATGADY----- 417
 QY 244 DRAEMLPBLGLTRESYEVHAFVPEMAIDLEAEKRLREVMDKVCV-----NIFVEGA 298
 DB 418 -----LFIPEFRAVPHGRMODDEKVEVCORHRSKGRNNITIVAGA 457
 QY 239 -----GVEAIVAEQAKQGEVPRDAFGHKL--DAVNPCKWFGEGEPQOMIAEK 345
 DB 458 LDDQNPVTANDVVDALIEL--GLDTVTTLGHVQGGTRVADRLAT--LQGVDAVK 512
 QY 346 TLVQSGTFAPASASN-----VDNRLLIKSCADLAWECAFRREGVIGHDS 391
 DB 513 AVL-----FTPEYSPILIGILENKIIMPLVESVKTLSVA-----TALENKDF 557
 QY 392 DNGNVLRAIEF 402
 DB 558 DKALSLRDER 568

RESULT 15
 US-09-815-242-13785
 Sequence 13785, Application US/09815242
 Patent No. US2002061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyckind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: BLITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815.242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for windows version 4.0
 SEQ ID NO 13785
 LENGTH: 320
 TYPE: PR
 ORGANISM: Salmonella typhi
 US-09-815-242-13785

Query Match 10.5%; Score 240.5; DB 9; Length 320;
 Best Local Similarity 24.8%; Pred. No. 1e-15;
 Matches 102; Conservative 59; Mismatches 122; Indels 129; Gaps 18;

QY 33 KKKVALITAGGALPCNSAIGSLIERYTEIDPSIEIICRGYKGL-----LLGDSYPTA 87
 DB 3 KKGVLITSGDALPQNMALIRGVYR--AALTSELEVMGTYDGLIGYEDRMVQLDRYSVD 60
 QY 88 EVRKAGVLCRFGSSVIGNSRVKLTNVKDCVKGGLVKEGDDPQKVAADQLVKDVILHT 147
 DB 61 MNR-----GGTFLGSAR--PPEPRD-----ENIRAVAIENLKKRGIDALIV 100
 QY 146 ICGDDTNTAADIAAFLARNYGLTVIGLPTVNDVPEPIKOSIGAMTAEGGAFYFVAVV 207
 DB 101 ICGDGSYKAKRL-----TEMGFPCIGLPGITIDNDIGTDTYTTIGYFALGTVVEALDRL 154
 QY 208 VANNANPRMLIVHEVGRNCGMLTAAVTAQERYKLDPRAEWL--PELGLTRESYEVHAFV 265
 DB 155 RDTSSSHORISIV--EVMGRYCGDLTAA-----IAGCEFIIVPEVEFNE----- 200
 QY 266 VPEMAIDLEAEK-----RLREVMDKVCNIPSEAGVE--AIVAEQAK 310
 DB 201 -----DLVAELIKGIANKKHAIVALTENHCVDDELAFTEKGTGETATVIGHQORG 254
 QY 311 GQEVPRDAFGHKLDAVNPCKWFGEGEPQOMIAEKTIYQKSGYFAPASASNVDDMLIKS 370
 DB 255 GSPVYD-----RLASRGAVA----- 272
 QY 371 CADLAVE-----CAFRREGVIGHDENGNVLRAIEPPRIKGGKPEPNIIDTW 417
 DB 273 -IDLLEHGGRGCGVIGNEQLVHED-----IIDALENMK--RPFK--SDW 312

Search completed: January 29, 2004, 14:43:34
 Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 15:54:19 : Search time 26.8953 Seconds
(without alignments)
3044.503 Million cell updates/sec

3044.503 Million cell updates/sec

Title: US-09-941-947a-8
Perfect score: 1985
Sequence: 1 MKGICILGATGSGIVSTUDV.....LQADQDAREVARDIKTLVA 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1985	100.0	394	10	US-09-934-903-4 Sequence 4, Appl1
2	1985	100.0	394	10	US-09-934-868-4 Sequence 64, Appl1
3	1985	100.0	394	11	US-09-941-947a-8 Sequence 8, Appl1
4	1215.5	61.2	396	12	US-10-381-779-117 Sequence 117, App
5	1052.5	53.0	394	12	US-10-381-779-119 Sequence 119, App
6	1045	52.6	398	12	US-10-381-779-100 Sequence 100, App
7	1045	52.6	398	12	US-10-381-779-118 Sequence 118, App
8	1042	52.5	405	12	US-10-381-779-122 Sequence 122, App
9	1007	50.7	397	12	US-10-381-779-99 Sequence 99, Appl1
10	1007	50.7	397	12	US-10-381-779-120 Sequence 120, App
11	999.5	50.4	388	12	US-10-381-779-101 Sequence 101, App
12	999.5	50.4	388	12	US-10-381-779-116 Sequence 116, App
13	994.5	50.1	386	12	US-10-381-779-97 Sequence 97, Appl1
14	892.5	45.0	394	12	US-10-381-779-103 Sequence 103, App
15	892.5	45.0	394	12	US-10-381-779-121 Sequence 121, App

15	886	44.6	402	12	US-10-381-779-102	Sequence 102, App
17	886	44.6	402	12	US-10-381-779-123	Sequence 123, App
18	818.5	41.2	477	9	US-09-923-556-6	Sequence 6, Appl1
19	818.5	41.2	477	10	US-09-967-025-2	Sequence 2, Appl1
20	818.5	41.2	477	12	US-10-381-779-128	Sequence 128, App
21	817.5	41.2	477	14	US-10-047-412A-10	Sequence 10, Appl1
22	814.5	41.0	477	9	US-09-923-556-2	Sequence 2, Appl1
23	813.5	41.0	477	14	US-10-047-412A-29	Sequence 29, Appl1
24	798.5	40.2	388	12	US-10-381-779-98	Sequence 98, Appl1
25	798.5	40.2	388	12	US-10-381-779-125	Sequence 125, App
26	731.5	36.9	385	12	US-10-18-713A-4	Sequence 4, Appl1
27	729.5	36.8	486	12	US-10-381-779-131	Sequence 131, App
28	721.5	36.3	487	12	US-10-259-194A-268	Sequence 268, App
29	718	36.2	386	12	US-10-381-779-124	Sequence 124, App
30	697	35.1	394	15	US-10-156-761-10100	Sequence 10100, A
31	692	34.9	392	10	US-09-738-626-57709	Sequence 5709, App
32	691	34.8	436	10	US-09-712-363-258	Sequence 258, App
33	691	34.8	436	12	US-10-381-779-104	Sequence 104, App
34	691	34.8	436	12	US-10-381-779-127	Sequence 127, App
35	681	34.3	406	12	US-10-381-779-126	Sequence 126, App
36	620	31.2	340	12	US-10-289-762-362	Sequence 362, App
37	596	30.0	1588	12	US-09-925-778-2	Sequence 2, Appl1
38	486	24.5	356	12	US-10-381-779-130	Sequence 130, App
39	473	23.8	368	12	US-09-882-227-500	Sequence 500, App
40	428.5	21.6	177	12	US-10-032-395-1	Sequence 1, Appl1
41	251	12.6	98	12	US-09-864-408A-8646	Sequence 8646, App
42	114	5.7	425	12	US-10-369-493-21773	Sequence 21773, A
43	113.5	5.7	1073	15	US-10-156-761-10416	Sequence 10416, A
44	109.5	5.5	6146	15	US-10-156-761-10416	Sequence 10416, A
45	109	5.5	426	12	US-10-369-493-21804	Sequence 21804, A

ALIGNMENTS

RESULT 1
US-09-934-903-4
Sequence 4, Application US/09934903
Patent No. US20020102690A1
GENERAL INFORMATION:
APPLICANT: Kofas, Matheos
APPLICANT: Odem, J. Martin
APPLICANT: Schenzle, Andreas J.
APPLICANT: No. US20020102690A1ton, Kelley C.
APPLICANT: Tomb, Jean-Francois
APPLICANT: Rouviere, Pierre
APPLICANT: Picataggio, Stephen
APPLICANT: Cheng, Qiong
TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
FILE REFERENCE: C1646 US NA
CURRENT APPLICATION NUMBER: US/09/934,903
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,907
PRIOR FILING DATE: September 1, 2001
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 4
LENGTH: 394
TYPE: PRT
ORGANISM: Methylobionas 16a
FEATURE:
OTHER INFORMATION: Amino acid sequences encoded by ORF2
US-09-934-903-4

Query Match 100.0%, Score 1985, DB 10, Length 394;
Best Local Similarity 100.0%, Pred. No. 7.2e-186;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKGICILGATGSGIVSTUDVVAHHPDKQVVALTANGNIDALYECIAHREYAVVWES 60
DB 1 MKGICILGATGSGIVSTUDVVAHHPDKQVVALTANGNIDALYECIAHREYAVVWES 60
QY 61 KVAERKQRLASPVADIVLSSSEALQGVATLENVDVYMAALVGAAGLPTLAAKAGKT 120

```

Db      61 KVAEFKORIAASPVADIKVLSGSEALQOVATLENVDVMAAIVGAAGLFTLLAAARAGKT 120
Qy      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Db      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Qy      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Db      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Qy      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Db      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Qy      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Db      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Qy      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394
Db      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394

```

RESULT 2

```

US-09-934-868-64
; Sequence 64, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M.
; APPLICANT: Schenaz, Andreas J.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL4596 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 64
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; OTHER INFORMATION: Amino acid sequences encoded by DXR
US-09-934-868-64

```

```

Query Match      100.0%; Score 1985; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 7, 2e-186;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MKGICILGATSGISVSTLDVVAARHPDKYQVVALTANENIDALYEQCLAHHEPVAVVMES 60
Db      1 MKGICILGATSGISVSTLDVVAARHPDKYQVVALTANENIDALYEQCLAHHEPVAVVMES 60
Qy      61 KVAEFKORIAASPVADIKVLSGSEALQOVATLENVDVMAAIVGAAGLFTLLAAARAGKT 120
Db      61 KVAEFKORIAASPVADIKVLSGSEALQOVATLENVDVMAAIVGAAGLFTLLAAARAGKT 120
Qy      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Db      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Qy      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Db      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Qy      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Db      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Qy      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Db      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360

```

```

Db      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Qy      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394
Db      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394

```

RESULT 3

```

US-09-941-947a-8
; Sequence 8, Application US/09941947a
; Publication No. US2003000528A1
; GENERAL INFORMATION:
; APPLICANT: Bzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL4903 US NA
; CURRENT APPLICATION NUMBER: US/09/941, 947a
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229, 907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229, 858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Methylomonas 16a
US-09-941-947a-8

```

```

Query Match      100.0%; Score 1985; DB 11; Length 394;
Best Local Similarity 100.0%; Pred. No. 7, 2e-186;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MKGICILGATSGISVSTLDVVAARHPDKYQVVALTANENIDALYEQCLAHHEPVAVVMES 60
Db      1 MKGICILGATSGISVSTLDVVAARHPDKYQVVALTANENIDALYEQCLAHHEPVAVVMES 60
Qy      61 KVAEFKORIAASPVADIKVLSGSEALQOVATLENVDVMAAIVGAAGLFTLLAAARAGKT 120
Db      61 KVAEFKORIAASPVADIKVLSGSEALQOVATLENVDVMAAIVGAAGLFTLLAAARAGKT 120
Qy      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Db      121 VLANKKALVMSGGIFMQAVSDSGAVLLPIDSSENAIFQCPAGYTPGHTAKARRILLT 180
Qy      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Db      181 ASGGFRRTPIETLSSTVPDOAVAHPRKDMGRKISVDSATMMKGLIELACLLFNNRBD 240
Qy      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Db      241 QIEVVIHQSIHSMVVDVDSVLAQKGNPDMRTPIAHAMAMPERFDSGVAPLDIEVGH 300
Qy      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Db      301 MDEFKPDIKRPPCRLAYEALIKSGGIMPTVLANAETIAVAFLNEEVEKFTDIAVIERSM 360
Qy      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394
Db      361 AOFKPDAGSLEVLQADQDAREVARDIITLVA 394

```

```

RESULT 4
US-10-381-779-117
; Sequence 117, Application US/10381779

```

```

; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-381-779-117

Query Match      61.2%; Score 1215.5; DB 12; Length 396;
Best Local Similarity 64.1%; Pred. No. 2.1e-110;
Matches 245; Conservative 48; Mismatches 88; Indels 1; Gaps 1;

QY 4 ICILGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 ISVIGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 EFKQRIASPVDIVLVSSEALQOVATLENVDVTMAIYGAAGLPTLAAAXGKTVL 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 ALQSLAAGR-TRVLFGEQALCEVASAPEVDVMAIYGAAGLPTLAAAXGKTVL 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 124 ANKEALVMSGOIFMQAVSDSGAVLLPIDSEHNAIPQCMFAGYTGHTAKQARRILLTASG 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 ANKEALVMSGALFMQAVKSGAVLLPIDSEHNAIPQCLPRVYADGLERGVRRILLTASG 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 184 GPFRTPIETLSSVTPDOAVAHKPMWGRKISVDSATMMNGLELIEAHWLFNCPCPDL 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 GPFRTPIETLSSVTPDOAVAHKPMWGRKISVDSATMMNGLELIEAHWLFNCPCPDL 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 244 VVIHPOSIIHSMVDVYDGSVLAQMGNDPMTPIAHAMWPERFDSGVAPLDIFEVGHMDF 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 VVIHPOSIIHSMVDVYDGSVLAQMGNDPMTPIAHAMWPERFDSGVAPLDIFEVGHMDF 305
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 304 EKELKRPFCRLAIEAIAKSGIMPTVLANAETAVAEFLNEEVKFTDIATITERSMAQF 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 QRPDEQRPFCRLAIAKSGIMPTVLANAETAVAEFLNEEVKFTDIATITERSMAQF 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 364 KPDAGSLELVLAQDODAREYVA 385
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 AVTAVESLDQVLAADRRARSA 387
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-10-381-779-119
; Sequence 119, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; SBO ID NO 119

```

```

; LENGTH: 394
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-381-779-119

Query Match      53.0%; Score 1052.5; DB 12; Length 394;
Best Local Similarity 55.5%; Pred. No. 2.1e-94;
Matches 216; Conservative 49; Mismatches 123; Indels 1; Gaps 1;

QY 4 ICILGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 LTIIIGTSGISGISTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 EFKQRIASPVDIVLVSSEALQOVATLENVDVTMAIYGAAGLPTLAAAXGKTVL 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 RLEALLKEDGTR-TQVLTGAQALVDVASADEVSGVCAIVAGVLPRLAALAAQKQKTVL 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 124 ANKEALVMSGOIFMQAVSDSGAVLLPIDSEHNAIPQCMFAGYTGHTAKQARRILLTASG 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 ANKEALVMSGALFMQAVKSGAVLLPIDSEHNAIPQCLPRVYADGLERGVRRILLTASG 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 184 GPFRTPIETLSSVTPDOAVAHKPMWGRKISVDSATMMNGLELIEAHWLFNCPCPDL 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 GPFRTPIETLSSVTPDOAVAHKPMWGRKISVDSATMMNGLELIEAHWLFNCPCPDL 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 244 VVIHPOSIIHSMVDVYDGSVLAQMGNDPMTPIAHAMWPERFDSGVAPLDIFEVGHMDF 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 VVIHPOSIIHSMVDVYDGSVLAQMGNDPMTPIAHAMWPERFDSGVAPLDIFEVGHMDF 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 304 EKELKRPFCRLAIEAIAKSGIMPTVLANAETAVAEFLNEEVKFTDIATITERSMAQF 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 QRPDEQRPFCRLAIEAIAKSGIMPTVLANAETAVAEFLNEEVKFTDIATITERSMAQF 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 364 KPDAGSLELVLAQDODAREYVA 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 FSDGIGDIGVLAQDARRARARAFITGL 393
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-10-381-779-100
; Sequence 100, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jensen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-381-779-100

Query Match      52.6%; Score 1045; DB 12; Length 398;
Best Local Similarity 54.9%; Pred. No. 1.1e-93;
Matches 219; Conservative 56; Mismatches 118; Indels 6; Gaps 3;

QY 1 MKGICIGATGSIQVSTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKQITLIGTSGISGISTLDVVAARHPDKYQVVALTANGNIDALYEQCLAHHPERYAVVVMESKVA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 KVAEFKQRIASPVDIVLVSSEALQOVATLENVDVTMAIYGAAGLPTLAAAXGKTVL 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 S-AKGLKTMKQQGGSRTVLSGGQAAQMAALEDVDQVMAIYGAAGLPTLAAAXGKTVL 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 121 VILANKKALVMSGOIFMOAVSDSGAVLLPIIDSEHNAIFOCM--PAGYTPGHTAKAR--- 175
DB 120 IILANKESLVTCORLMDAVKQSKAGLLFVDSSEHNAIFQSLPQPIQHNLGVADLEQNGV 179
QY 176 RILITASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIACLLF 235
DB 180 SIILITSGGPFRETPRLATMTPTDQACHRNPMWSGRKISVDSATMANKGELYIARWL 239
QY 236 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDPMTPIAHAMAMPERFDSGVAPLDI 295
DB 240 NASASQMEVLHPOSIVHSVTRQODSVLAQLOEPKMTPIAHAMAMPERVNSGVKPLDF 299
QY 296 FEVGHMDFEKDPDKRPPCLRLAYEAIRKSGIMFTVLANANEIYAEALNEEYKFTDIAT 355
DB 300 CKLSALTFPAADRYPCCLAMEAEFGQATTAALNANEITVAALFQAQIRFTDIAL 359
QY 356 IERSMAQFKPDAGSLELVLAQDQAREVARDIITVA 394
DB 360 NLSVLERMDREBQCVDDVLSVDANAREVAREVRLAS 398

RESULT 7
US-10-381-779-118
; Sequence 118, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jeseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/0020US1
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-381-779-118

Query Match 52.6%; Score 1045; DB 12; Length 398;
Best Local Similarity 54.9%; Pred. No. 1,1e-93;
Matches 219; Conservative 56; Mismatches 118; Indels 6; Gaps 3;

QY 1 MKGICIGATSGISVSTLDVVAHPRKYVVALTANGNDIALYEQCLAHHPREYAVVWMS 60
DB 1 MKQITLISGTSIGSTLIDVVRNHEFRVALVAGNVTRMVEQCLERPRAVMDDEA 60
QY 61 KVAEFKQRIASPVADIKVLSGSEALQOVATLENVTVMATYGAAGLLPTLAAAGKT 120
DB 61 S-AKLITMLQOQSRTREVLSGQOACDMALBVDVMAIYGAAGLLPTLAAIAGKT 119
QY 121 VILANKKALVMSGOIFMOAVSDSGAVLLPIIDSEHNAIFOCM--PAGYTPGHTAKAR--- 175
DB 120 IILANKESLVTCORLMDAVKQSKAGLLFVDSSEHNAIFQSLPQPIQHNLGVADLEQNGV 179
QY 176 RILITASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIACLLF 235
DB 180 SIILITSGGPFRETPRLATMTPTDQACHRNPMWSGRKISVDSATMANKGELYIARWL 239
QY 236 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDPMTPIAHAMAMPERFDSGVAPLDI 295
DB 240 NASASQMEVLHPOSIVHSVTRQODSVLAQLOEPKMTPIAHAMAMPERVNSGVKPLDF 299
QY 296 FEVGHMDFEKDPDKRPPCLRLAYEAIRKSGIMFTVLANANEIYAEALNEEYKFTDIAT 355
DB 300 CKLSALTFPAADRYPCCLAMEAEFGQATTAALNANEITVAALFQAQIRFTDIAL 359

```

```

QY 356 IERSMAQFKPDAGSLELVLAQDQAREVARDIITVA 394
DB 360 NLSVLERMDREBQCVDDVLSVDANAREVAREVRLAS 398

RESULT 8
US-10-381-779-122
; Sequence 122, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jeseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/0020US1
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-381-779-122

Query Match 52.5%; Score 1042; DB 12; Length 405;
Best Local Similarity 53.2%; Pred. No. 2,3e-93;
Matches 210; Conservative 69; Mismatches 109; Indels 8; Gaps 3;

QY 1 MKGICIGATSGISVSTLDVVAHPRKYVVALTANGNDIALYEQCLAHHPREYAVVWMS 60
DB 7 MKQITLISGTSIGSTLIDVVRNHEFRVALVAGNVTRMVEQCLERPRAVMDDEA 66
QY 61 KVAEFKQRIASPVADIKVLSGSEALQOVATLENVTVMATYGAAGLLPTLAAAGKT 120
DB 67 AKKLAELKRNH-QSTTVLAQQAICELAHPEALMVAIYGAAGLLPTLSAVAGRR 125
QY 121 VILANKKALVMSGOIFMOAVSDSGAVLLPIIDSEHNAIFOCM-----AGYTPGHTAKA 174
DB 126 VILANKKALVTCGLFTDAVRESOQQLPVDSEHNAIFQSLPEARQIGFC-LSBLGI 184
QY 175 RILITASGGPFRRTPIETLSVTPDQAVAHPRKMDGRKISVDSATMANKGELIACLL 234
DB 185 SKIVITSGGPFRTPIEBQEQITTPQAVAHPRKMDGRKISVDSATMANKGELYIARWL 244
QY 235 NMEPDQLEVIHPOSIHSMVDYVDSVLAQMGNDPMTPIAHAMAMPERFDSGVAPLDI 294
DB 245 FNAASEEVLHPOSIVHSVTRQODSVLAQLOEPKMTPIAHAMAMPERVNSGVKPLDF 304
QY 295 IFEVGHMDFEKDPDKRPPCLRLAYEAIRKSGIMFTVLANANEIYAEALNEEYKFTDIAT 354
DB 305 FYQANGLTFIPDQIRPCLKLAIDAFSAQVATTAALNANEITVAALFQAQIRFTDIAR 364
QY 355 IERSMAQFKPDAGSLELVLAQDQAREVARDIITVA 389
DB 365 INQIVSVLQPKIHICIEDVLEVDKARRELSOSII 399

RESULT 9
US-10-381-779-99
; Sequence 99, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jeseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/0020US1

```

```

; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 397
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-10-381-779-99

```

```

Query Match      50.7% Score 1007; DB 12; Length 397;
Best Local Similarity 50.6%; Pred. No. 6e-90;
Matches 200; Conservative 76; Mismatches 111; Indels 8; Gaps 3;

```

```

QY 2 KGCITLGGATGSGVSTLTVARHPKQVVALTANGNIDALYEQCLAHPEYAVVWESK 61
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 QNIVLISGSGIGKSTLTVIENNPQKHAFLVGGKVEAMFECIKFRPHFALDDVNA 63
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 VAEFKORIAASPVADIKVLSGSEALQOVATLENVDTWMAAIVGAAGLLPTLAAAKAGT 121
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 64 AKILREKLIANH1-PTVYLAGRAICELAAHPDADQIMASIVGAAGLLPTLSAVAGKRV 122
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 LLAKEALVMSGOITMVAUSDGAVLLPIDEENHAIPOCF-----AGTTPGHAKQAR 175
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 123 LLANKESLVTCGQLFIDAVKNVGSLLPVDSEHNAIPQSIPPEAKKIGFCP-LSBLGVS 181
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 RILTLASGPPFRPTIETLSVTPDOAVAHPKMDGKISVDSATMNNKGLIELIACILF 235
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 182 KILTGSGGPPRYTLEQPTNITPQAVAHNWSGKKISVDSATMNNKGLIELIACILF 241
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 236 NMEPDQIEVVIHPQSIHSMWDYVDGSLAQMGNPDMPTPIAHMAWPERDSGVAPLDI 295
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 242 NASAEEMEVIIHPQSIHSMWRYVDSVITQMGNDMPPTIAETMAVPHRTFAGVEPLDF 301
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 296 FEVGMDEPEKPDLPKPCPLRLAYRAIKSGIMPTVLAANEIAVEAFINBEVFTDIAYI 355
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 302 FKIXELTIEPDNPNYPVLKALIDPAAGQVATTAMNANIEAIVQAFIDRQIGFMDIAXI 361
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 356 IERSVAQRPDDAGSLIYLQADDPAREVARDIIR 390
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 362 NSKITERISPYTIQINIDVLEIDAQAREIAKTLLR 396
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 10
US-10-381-779-120
; Sequence 120, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jesseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 397
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-10-381-779-120

```

```

Query Match      50.7% Score 1007; DB 12; Length 397;
Best Local Similarity 50.6%; Pred. No. 6e-90;

```

```

Matches 200; Conservative 76; Mismatches 111; Indels 8; Gaps 3;

```

```

QY 2 KGCITLGGATGSGVSTLTVARHPKQVVALTANGNIDALYEQCLAHPEYAVVWESK 61
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 QNIVLISGSGIGKSTLTVIENNPQKHAFLVGGKVEAMFECIKFRPHFALDDVNA 63
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 VAEFKORIAASPVADIKVLSGSEALQOVATLENVDTWMAAIVGAAGLLPTLAAAKAGT 121
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 64 AKILREKLIANH1-PTVYLAGRAICELAAHPDADQIMASIVGAAGLLPTLSAVAGKRV 122
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 LLAKEALVMSGOITMVAUSDGAVLLPIDEENHAIPOCF-----AGTTPGHAKQAR 175
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 123 LLANKESLVTCGQLFIDAVKNVGSLLPVDSEHNAIPQSIPPEAKKIGFCP-LSBLGVS 181
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 RILTLASGPPFRPTIETLSVTPDOAVAHPKMDGKISVDSATMNNKGLIELIACILF 235
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 182 KILTGSGGPPRYTLEQPTNITPQAVAHNWSGKKISVDSATMNNKGLIELIACILF 241
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 236 NMEPDQIEVVIHPQSIHSMWDYVDGSLAQMGNPDMPTPIAHMAWPERDSGVAPLDI 295
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 242 NASAEEMEVIIHPQSIHSMWRYVDSVITQMGNDMPPTIAETMAVPHRTFAGVEPLDF 301
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 296 FEVGMDEPEKPDLPKPCPLRLAYRAIKSGIMPTVLAANEIAVEAFINBEVFTDIAYI 355
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 302 FKIXELTIEPDNPNYPVLKALIDPAAGQVATTAMNANIEAIVQAFIDRQIGFMDIAXI 361
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 356 IERSVAQRPDDAGSLIYLQADDPAREVARDIIR 390
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 362 NSKITERISPYTIQINIDVLEIDAQAREIAKTLLR 396
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 11
US-10-381-779-101
; Sequence 101, Application US/10381779
; Publication No. US20030219798A1
; GENERAL INFORMATION:
; APPLICANT: Gokarn, Ravi R
; APPLICANT: Jesseen, Holly
; APPLICANT: Zidwick, Mary Jo
; TITLE OF INVENTION: Isoprenoid Production
; FILE REFERENCE: 12904/002US1
; CURRENT APPLICATION NUMBER: US/10/381,779
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30328
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,580
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 388
; TYPE: PRF
; ORGANISM: Zymomonas mobilis
US-10-381-779-101

```

```

Query Match      50.4% Score 999.5; DB 12; Length 388;
Best Local Similarity 50.9%; Pred. No. 3.2e-89;
Matches 199; Conservative 74; Mismatches 109; Indels 9; Gaps 3;

```

```

QY 4 ICIGATGSGVSTLTVARHPKQVVALTANGNIDALYEQCLAHPEYAVVWESKVA 63
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 7 TVIGATGSGVSTLTVARHPKQVVALTANGNIDALYEQCLAHPEYAVVWESKVA 66
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 EFKORIAASPVADIKVLSGSEALQOVATLENVDTWMAAIVGAAGLLPTLAAAKAGT 123
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 67 DLKELAGS---SVEAAGADALVEAMM-GADVTMAAIIIGAGLKATLAIRKGTVAL 122
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 124 ANKEALVMSGOITMVAUSDGAVLLPIDEENHAIPOCFMPCGYPTGHAKQAR 183
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 123 ANKESLVASGGLMDAAREHGTLLPVDSEHNAIPQCF-----PHNRDYVRKILITASG 177
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 184 GPFRTPTETLSVTPDOAVAHPKMDGKISVDSATMNNKGLIELIACILF 243
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 178 GFFRTSLAEMATVTPERAIVQHEHNSMGAKISIDSATMKNKLELEAVHLFOIPEKFE 237
 Qy 244 VVIHPOSIHSMVDYVDSVLAQMGNDPKRTPIAHMAMPERFDGVAFLDIEVGEKDF 303
 Db 238 IIVHPQSVIHSWVEYIDGSIILQISGPDKRTPIGHTLAMPKEMETPASLDPTKROKDF 297
 Qy 304 EKEDLKRFPCILRLAYEAIKSGGIMPTVLAANEIYAEALNEEVEKFTDIIVIERMAQF 363
 Db 298 EADYERFPALTLAMEKISGSGAPVMAANEIYAAFLDKKIGFLDIKTIKTEKLDHY 357
 Qy 364 KPDDAGSLLEVLQADQAREVARDIKTIYA 394
 Db 358 TPATPSLEDPVPAIDNEARQAAALMESLPA 388

RESULT 12
 US-10-381-779-116
 ; Sequence 116, Application US/10381779
 ; Publication No. US20030219798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gokarn, Ravi R
 ; APPLICANT: Jessen, Holly
 ; APPLICANT: Zidwick, Mary Jo
 ; TITLE OF INVENTION: Isoprenoid Production
 ; FILE REFERENCE: 12904/002US1
 ; CURRENT APPLICATION NUMBER: US/10/381,779
 ; CURRENT FILING DATE: 2003-03-28
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: 60/236,580
 ; NUMBER OF SEQ ID NOS: 190
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 116
 ; LENGTH: 388
 ; TYPE: PRT
 ; ORGANISM: Zymomonas mobilis
 US-10-381-779-116

Query Match 50.4%; Score 999.5; DB 12; Length 388;
 Best Local Similarity 50.9%; Pred. No. 3.2e-89;
 Matches 199; Conservative 74; Mismatches 109; Indels 9; Gaps 3;
 Qy 4 ICILGATGSIQVSTLDVVAHHPDKYQVVALTANGNIDALYEQCLAHHPYAVVMESKVA 63
 Db 7 VTVIGATGSIQVSTLDLIERLDRYQVALTANGNVMDLADAAKTVAKAVIADPSLVN 66
 Qy 64 EFKORIAAPVADIKVLSGSEALQOVATLENDVTWMAIYGAAGLPTLAAAKAGKTVLL 123
 Db 67 DKKEALGSG--SEVAAAGADALVEAMM--GADWTVAALIGCAELKATTLAIAIRKGTVAL 122
 Qy 124 ANKEALVNSGQIFQAVSDSGAVLLPTDSEHNAIFQCPKAGYTPGHTAXOARRILLTASG 183
 Db 123 ANKEALVNSGQIFQAVSDSGAVLLPTDSEHNAIFQCPKAGYTPGHTAXOARRILLTASG 177
 Qy 184 GFFRTPIETLSVTPDQAVAHPKWDMGRKISVDSATMKNKLELEAVHLFOIPEKFE 243
 Db 178 GFFRTSLAEMATVTPERAIVQHEHNSMGAKISIDSATMKNKLELEAVHLFOIPEKFE 237
 Qy 244 VVIHPOSIHSMVDYVDSVLAQMGNDPKRTPIAHMAMPERFDGVAFLDIEVGEKDF 303
 Db 238 IIVHPQSVIHSWVEYIDGSIILQISGPDKRTPIGHTLAMPKEMETPASLDPTKROKDF 297
 Qy 304 EKEDLKRFPCILRLAYEAIKSGGIMPTVLAANEIYAEALNEEVEKFTDIIVIERMAQF 363
 Db 298 EADYERFPALTLAMEKISGSGAPVMAANEIYAAFLDKKIGFLDIKTIKTEKLDHY 357
 Qy 364 KPDDAGSLLEVLQADQAREVARDIKTIYA 394
 Db 358 TPATPSLEDPVPAIDNEARQAAALMESLPA 388

RESULT 13

US-10-381-779-97
 ; Sequence 97, Application US/10381779
 ; Publication No. US20030219798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gokarn, Ravi R
 ; APPLICANT: Jessen, Holly
 ; APPLICANT: Zidwick, Mary Jo
 ; TITLE OF INVENTION: Isoprenoid Production
 ; FILE REFERENCE: 12904/002US1
 ; CURRENT APPLICATION NUMBER: US/10/381,779
 ; CURRENT FILING DATE: 2003-03-28
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: 60/236,580
 ; NUMBER OF SEQ ID NOS: 190
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 97
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: Spilngomonas truegeri
 US-10-381-779-97

Query Match 50.1%; Score 994.5; DB 12; Length 386;
 Best Local Similarity 51.0%; Pred. No. 9.7e-89;
 Matches 203; Conservative 69; Mismatches 109; Indels 17; Gaps 4;
 Qy 1 MKGICIGATGSIQVSTLDVVAHHPDKYQVVALTANGNIDALYEQCLAHHPYAVVMES 60
 Db 2 VRTVTVIGATGSIQVSTLDLIERLDRYQVALTANGNVMDLADAAKTVAKAVIADPSLVN 61
 Qy 61 KVAEFKORIAAPVADIKVLSGSEALQOVATLENDVTWMAIYGAAGLPTLAAAKAGKTV 120
 Db 62 CLPALQERLAGS--GVEAMGAGHSYCDVARM--GADWTVAALIGSGELGTVAAALAGGT 117
 Qy 121 VLAANKALVNSGQIFQAVSDSGAVLLPTDSEHNAIFQCPKAGYTPGHTAXOARRILLTASG 176
 Db 118 VALANKALVNSGQIFQAVSDSGAVLLPTDSEHNAIFQCPKAGYTPGHTAXOARRILLTASG 168
 Qy 177 ILTLASGQFFRTPIETLSVTPDQAVAHHPDKWDMGRKISVDSATMKNKLELEAVHLFOI 226
 Db 169 IILTLASGQFFRTPIETLSVTPDQAVAHHPDKWDMGRKISVDSATMKNKLELEAVHLFOI 228
 Qy 227 MEPOQIEVVIHPOSIHSMVDYVDSVLAQMGNDPKRTPIAHMAMPERFDGVAFLDIEV 296
 Db 229 VAAQQLAVLHRSQVHSWVEYIDGSIILQISGPDKRTPIGHTLAMPKEMETPASLDPTK 288
 Qy 297 EYGHMDFEKPDLKRFPCILRLAYEAIKSGGIMPTVLAANEIYAEALNEEVEKFTDIIVIER 356
 Db 289 TVGKLEFENPDLDFPALALAMEALKXGAPATLUNANEIYAAFLAGRIIGFLDIKTIK 348
 Qy 357 ERSMAQFEPDDAGSLLEVLQADQAREVARDIKTIYA 394
 Db 349 ADTLERYDPAPBTLDVLAIDAEARLYAARVQCVAA 386

RESULT 14
 US-10-381-779-103
 ; Sequence 103, Application US/10381779
 ; Publication No. US20030219798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gokarn, Ravi R
 ; APPLICANT: Jessen, Holly
 ; APPLICANT: Zidwick, Mary Jo
 ; TITLE OF INVENTION: Isoprenoid Production
 ; FILE REFERENCE: 12904/002US1
 ; CURRENT APPLICATION NUMBER: US/10/381,779
 ; CURRENT FILING DATE: 2003-03-28
 ; PRIOR APPLICATION NUMBER: PCT/US01/30328
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: 60/236,580
 ; NUMBER OF SEQ ID NOS: 190

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 103
; LENGTH: 394
; TYPE: prt
; ORGANISM: Synechocystis sp. PCC 6803
US-10-381-779-103

```

Query Match	45.0%;	Score 892.5;	DB 12;	Length 394;
Best Local Similarity	47.3%;	Pred. No. 1e-78;		
Matches 189;	Conservative 63;	Mismatches 128;	Indels 13;	Gaps 3;

[illegible]

```

RESULT 15
US-10-381-779-121
: Sequence 121. Application US/10381779
: Publication No. US20030219798A1
: GENERAL INFORMATION:
: APPLICANT: Gokarn, Ravi R
: APPLICANT: Jessen, Holly
: APPLICANT: Zidwick, Mary Jo
: TITLE OF INVENTION: Isoprenoid Production
: FILE REFERENCE: 12904/002US1
: CURRENT APPLICATION NUMBER: US/10/381, 779
: CURRENT FILING DATE: 2003-03-28
: PRIOR APPLICATION NUMBER: PCT/US01/30328
: PRIOR FILING DATE: 2001-09-28
: PRIOR APPLICATION NUMBER: 60/236,580
: PRIOR FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 190
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 121
: LENGTH: 394
: TYPE: PRT
: ORGANISM: Synechocystis sp. PCC 6803
US-10-381-779-121

```

Query Match	45.0%;	Score 892.5;	DB 12;	Length 394;
Best Local Similarity	47.3%;	Pred. No. 1e-78;		
Matches 183; Conservative	63;	Mismatches 128;	Indels 13;	Gaps 3;

0y _ MKGSCILAAAGSGVSTLVVVAHPKPKVVAALVNGVINDVLPQCLAAHPETAUVVMES 60
2 VKGISTLGGSGSGITPDIIVTHHPDAFVGVVGLAAGVALLAAQVVAFFREPIYALIQAE 61
0y 61 KVAFKQRIAAFPVADIKVLSSGEMLDQVATLENDVTMAAIVGAAGLLPTLAARAKGKT 120

[illegible]

Search completed: January 29, 2004, 16:21:15
Job time : 27.8953 secs